

SEARCH REQUEST FORM

50080

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: CS/Team. Tully
 Searcher Phone #: 308-45010
 Searcher Location: Biotec Lib
 Date Searcher Picked Up: 8/31/01
 Date Completed: 9/4/01
 Searcher Prep & Review Time: _____
 Clerical Prep Time: 4 min
 Online Time: 3 min

Type of Search

NA Sequence (#) _____
 AA Sequence (#) 1
 Structure (#) _____
 Bibliographic _____
 Litigation _____
 Fulltext _____
 Patent Family _____
 Other _____

Vendors and cost where applicable

STN _____
 Dialog _____
 Questel/Orbit _____
 Dr.Link _____
 Lexis/Nexis _____
 Sequence Systems ABSSO1
 WWW/Internet _____
 Other (specify) _____

STIC-Biotech/ChemLib

50080

From: Jiang, Dong
Sent: Thursday, August 30, 2001 6:15 PM
To: STIC-Biotech/ChemLib
Subject: SN09/333,159

6/14/99
McCarthy

Please search SEQ ID NO:47 (back translation only)

-issued
-commercial

Please send results on paper to Dong Jiang in 10B-01 (mail stop CM1-10C01). Thank you very much for your help.

Dong Jiang (78243)
703-305-1345
U.S. Patent and Trademark Office
Art Unit 1646
dong.jiang@uspto.gov
CM1-10B01
Mail stop: CM1-10C01

RECEIVED
AUG 31 2001
STIC

Percent Similarity: 83.171 Percent Identity: 60.732

alignment_block:

US-09-333-159-47 x HSLAL

Align seg 1/1 to: HSLAL from: 1 to: 2626

```

2 LeuGluThrLeuSerArgGlnTrpIleValSerHisArgMetGluMetTr 18
|||||
113 CTCGAGACAGGCGCCGCGAG.....GACAGCTCCAGAAATGAATAATGCG 156
|||||
18 pLeuLeuIleValAlaTrpMetPheGlnArgAsnValAsnSerValH 35
|||||
157 GTTCTGGGGTGGTGGTCTGTTGGTCTCTGGACCCCTGCATTCTGAGG 206
|||||
35 lsMetProThrLys.....AlaValAspProGluAlaPheMetAsnIle 49
|||||
207 GGTCTGGAGGGAACACTGACAGCTGTGGATCTCTGAACAAACATGAATGTG 256
|||||
50 SerGluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGluValAl 66
|||||
257 AGTGAATATATCTTACTGGGATTCCTAGTGAGGAATACCTAGTTGA 306
|||||
66 aThrGluAspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuV 83
|||||
307 GACAGAAGATGGATATATCTGTGCTTACCAGATTCCTCATGGGAGGA 356
|||||
83 alGlnProLysLysThrGlySerArgProValValLeuLeuGlnHisGly 99
|||||
357 AGAACCATCTTGACAAAGGTCACAAACAGTGTGCTTCTGCAACATGGC 406
|||||
100 LeuValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnSerLe 116
|||||
407 TTGCTGGCAGATTCTAGTAATCTGGGTGACAAACCTTGCCACAGCAGCCT 456
|||||
116 uClyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyAsnSerA 133
|||||
457 GGGCTTCACTTCTGCTGATGCTGTTGGTGGTGGTGGTGGTGGTGGTGG 506
|||||
133 rGlyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAspGlnAsp 149
|||||
507 GAGGAATACCTGCTCGGAAACATAAGACACTCTCAGTTCTCAGGAT 556
|||||
150 GluPheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeuProAl 166
|||||
557 GAATCTGGGCTTTCAGTTATGATGATGATGATGATGATGATGATGATGAT 606
|||||
166 aValIleAsnPheIleLeuGlnLysThrGlyGlnGluLysIleTyrTrv 183
|||||
607 TTCCATTAACTTCACTTGAATAAAGCTGCCAAGACAAAGTGTATTATG 656
|||||
183 alGlyTyrSerGlnGlyThrMetGlyPheIleAlaPheSerThrMet 199
|||||
657 TGGCTCATCTCAAGGACCACTATAGTGTATATAGCATTTTCACAGATC 706
|||||
200 ProGluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaProIleAl 216
|||||
707 CCTAGGTGGCTAAAGAGATTAATGTTTTCCTGGTGGTGGTGGTGGTGG 756
|||||
216 aThrValLysHisAlaLysSerProGlyThrLysPheLeuLeuLeuProA 233
|||||
757 TTCCGTGCGCTCTGTACTAGCCCTATGCGCAATATTAGGACGATTACCAG 806
|||||
233 spMetMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThr 249
|||||
807 ATCATCTCAATTAGGACTATTTTGGACAAAGAAATTTCTTCCCGAGT 856
|||||
250 ArgPheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAs 266
|||||
857 GCGTTTGAAGTGGCTGGGTACCCAGTGTGGCACTCATGCTCATCTACATG 906
|||||
266 pGlnIleCysSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnA 283
|||||

```

```

907 GGAGCTCTGTGGAAATCTCTGTTTCTCTGTGTGGATTTAATGAGAGAA 956
|||||
283 snMetAsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGly 299
|||||
957 ATTTAAATATGTCTAGAGTGGATGATATACACACATTTCTCTGCTGGA 1006
|||||
300 ThrSerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyG1 316
|||||
1007 ACTTCTGTGCAAAACATGTTACACTGGAGCCAGGCTGTTAAATTTCCAAA 1056
|||||
316 uLeuArgAlaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysA 333
|||||
1057 GTTTCAGCCCTTGGACTGGGAGCAGCTGCCAAGAATATTTTCATTACA 1106
|||||
333 snGlnProThrProValArgTyrArgValArgAspMetThrValProThr 349
|||||
1107 ACCAGATTATCTCTCCACATACAATGTGAAGGACATGCTTGTGCCGACT 1156
|||||
350 AlaMetTrpThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLy 366
|||||
1157 GCAGTCTGGAGCGGGGCTCAGACTGGCTTGCAGATCTCTACGACGTCAA 1206
|||||
366 smetLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsnIleProG 383
|||||
1207 TATCTTACTGACTCAGATCACCACACTTGGTGTTCATGAGAGCATTCGGG 1256
|||||
383 luTrpAaHisValAspPheIleTrpGlyLeuAspAlaProHisArgMet 399
|||||
1257 AATGGGAGCATCTGACTTCTATTTGGGCGCTGGATGCCCTTGGAGGCTT 1306
|||||
400 TyrAsnGluIleIleHisLeuMetGlnGln 409
|||||
1307 TATAATAAAATTTATTAATCTAATGAGGAAA 1336
|||||

```

seq_name: bp-pr-q: HUMLIPCHL

seq_documentation_block:

LOCUS HUMLIPCHL 2493 bp mRNA PRI 07-JAN-1995

DEFINITION Human lysosomal acid lipase/cholesterol esterase mRNA, complete cds.

ACCESSION M74775

VERSION M74775.1 GI:187151

KEYWORDS lysosomal acid lipase/cholesterol esterase.

SOURCE Homo sapiens fibroblast cDNA to mRNA.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2493)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Anderson, R.A. and Sando, G.N.

TITLE Cloning and expression of cDNA encoding human lysosomal acid

lipase/cholesterol ester hydrolase. Similarities to gastric and

lingual lipases

J. Biol. Chem. 266 (33), 22479-22484 (1991)

MEDLINE 92042192

FEATURES

Location/Qualifiers

1..2493

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_type="fibroblast"

/map="10"

<1..2493

/gene="LIPA"

/note="G00-120-153"

1..2493

/gene="LIPA"

41..1240

/gene="LIPA"

/EC_number="3.1.1.13"

/note="alternate initiator ATG may be at position 47"

/codon_start=1

/db_xref="GDB:G00-120-153"

/product="lysosomal acid lipase/cholesterol esterase"

/protein_id="AA59519.1"

CDS

gene

mRNA

925	CAGGCTGTAAATTC	AAAAGTTT	CAAGCCTTT	GACTGGGGAAGCAGTGC	974
326	rLysAsnLeuGluLys	CysAsnGlnProThr	ProValArgVala	343	
975	CAAGAAATATTTC	TATCAACAC	CAGAGTTATCCT	CCCACATCAATGTGA	1024
343	rgAspMetThrVal	ProThrAlaMet	TrpThrGlyGlyGlnAsp	TrpLeu	359
1025	AGGACATCCTGTG	CGCGACTGC	AGTCTGGAGCGGGGT	CACGACTGGCCT	1074
360	SerAsnProGluAsp	ValLysMetLeuLeuSer	GluValThrAsnLeu	11	376
1075	GCAGATGTCTACG	AGCTCAATATCTT	ACTGACTCAGAT	CACCAACTTGGT	1124
376	eTyH1sLysAsnIle	ProGluTrpAlaHis	ValAspPheIle	TrpGlyL	393
1125	GTTCCATGAGGAC	ATTCGGAGTTGG	AGCATCTTGACTT	CACTTTGGGGCC	1174
393	euAspAlaProHis	ArgMetTyAsnGlu	IleIleHisLeuMet	GlnGln	409
1175	TGGATGCCCTTGG	AGGCTTTATATAA	TAATTAATTAAT	TAATTAATGAGGAA	1224

seq_name: qb_pr10:HSU08464

seq_documentation_block: .
LOCUS HS008464 2481 bp mRNA PRI
DEFINITION Human lysosomal acid lipase mRNA, complete cds.

ACCESSION U08464
VERSION U08464.1 GI:505052

KEYWORDS
SOURCE . human.

ORGANISM	Homo sapiens	Fukushima
1	100	100
2	100	100
3	100	100
4	100	100
5	100	100
6	100	100
7	100	100
8	100	100
9	100	100
10	100	100
11	100	100
12	100	100
13	100	100
14	100	100
15	100	100
16	100	100
17	100	100
18	100	100
19	100	100
20	100	100
21	100	100
22	100	100
23	100	100
24	100	100
25	100	100
26	100	100
27	100	100
28	100	100
29	100	100
30	100	100
31	100	100
32	100	100
33	100	100
34	100	100
35	100	100
36	100	100
37	100	100
38	100	100
39	100	100
40	100	100
41	100	100
42	100	100
43	100	100
44	100	100
45	100	100
46	100	100
47	100	100
48	100	100
49	100	100
50	100	100
51	100	100
52	100	100
53	100	100
54	100	100
55	100	100
56	100	100
57	100	100
58	100	100
59	100	100
60	100	100
61	100	100
62	100	100
63	100	100
64	100	100
65	100	100
66	100	100
67	100	100
68	100	100
69	100	100
70	100	100
71	100	100
72	100	100
73	100	100
74	100	100
75	100	100
76	100	100
77	100	100
78	100	100
79	100	100
80	100	100
81	100	100
82	100	100
83	100	100
84	100	100
85	100	100
86	100	100
87	100	100
88	100	100
89	100	100
90	100	100
91	100	100
92	100	100
93	100	100
94	100	100
95	100	100
96	100	100
97	100	100
98	100	100
99	100	100
100	100	100

Mammal

AUTHORS Du, H.

and Hu

**JOURNAL
REFERENCE**

AUTHORS	TITLE
DU, H.	Direct

JOURNAL
Submit
childr

OH 452

source

E / EMP

sig_peptide

NOTES

mat_peptide

[illegible]


```

215 leAlaThrVallyshisAlaLysSerProGlyThrLysPheLeuLeuLeu 231
834 TGTTCCTCTCAATTTCTCGGGCCGGATGGTCAAAATGGGAGGTG 883
232 ProAspMetMetIleLysGlyLeuPheGlyLysGluPheLeuTyrG1 248
884 CCGGATCTGCTCTTGAGGACTTGTGGGCGAAGCAATTTCTCTCA 933
248 nThrArgPheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleL 265
934 GAGTCAATGGTGAATGGTAAGCACTACATTTGCACGACGTCATCA 983
265 euAspGlnIleCysSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThr 281
984 TGAAGGAGCTTGGCCCAACATCTCTTCTGATATGGTGGTCAATGAG 1033
282 AsnAsnMetAsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAl 298
1034 AAGATTTAAACATGCTAGAGTGGATGTGTATACACACACTGCTCTGC 1083
298 aclyThrSerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerG 315
1084 GGAACCTCTGTGCAAAACATGGTACACTGGACCCAGGTTGTTAAATACC 1133
315 lyGluLeuArgAlaPheAspTrpGlySerGluThrLysAsnLeuLys 331
1134 ATAAGCTTCAAGCTTTGACTGGGAAGCAGTGAAGAATATTATTCAT 1183
332 CysAsnGlnProThrProValArgTyrArgValArgAspMetThrValPr 348
1184 TACAATCAGATTATCCACCTTGTACAGCATAAAGACATGCAACTACC 1233
348 oThrAlaMetTrpThrGlyGlnAspTrpLeuSerAsnProGluAspV 365
1234 CACTGCTTTATGGAGCGGGGCAAGGACTGGTGGCAGACACCAGTGACA 1283
365 alLysMetLeuLeuSerGluValThrAsnLeuLeuIleTyrHisLysAsnIle 381
1284 TCAATATCTTACTGACTGAAATCCCACTGAGTGTAGTGTACCAAGAACAATT 1333
382 ProGluTrpAlaHisValAspPheIleTrpGlyLeuAspAlaProHisar 398
1334 CCCGAGTGGGACCATCTCGATTTATCTGGGGTTGGATGCCCCCTGGAG 1383
398 gMetTyrAsnGluIleIleHisLeuMetGlnGln 409
1384 GCTGTATAATGAAGTAGTCAGTCAATCAAGAA 1417

seq_name: 9b_rol:MMLYACLY

seq_documentation_block:
LOCUS MMLYACLY 2358 bp mRNA ROD 24-FEB-1999
DEFINITION M.musculus (C57 Black/6X CBA) LAL mRNA for lysosomal acid lipase.
ACCESSION 231689
VERSION 231689.1 GI:4456670
KEYWORDS LAL; lipase; lysosomal acid lipase.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2358)
AUTHORS Du,H., Witte,D.P. and Grabowski,G.A.
TITLE Tissue and cellular specific expression of murine lysosomal acid
lipase mRNA and protein
JOURNAL J. Lipid Res. 37 (5), 937-949 (1996)
MEDLINE 96363957
REFERENCE 2 (bases 1 to 2358)
AUTHORS Du,H.
TITLE Direct Submission
JOURNAL Submitted (05-APR-1994) Hong Du, Division of Human Genetics,
Children's Hospital Medical Center, 3333 Bernet Street, Cincinnati,
OH, 45229-3039, USA

```

FEATURES

```

source
1..2358
/organism="Mus musculus"
/strain="C57 Black/6X CBA"
/db_xref="taxon:10090"
/sex="Female"
/clone="mLAL #49"
/tissue_type="liver"
/clone_lib="Mouse liver cDNA library /Lamda-ZAP"
19..1212
/genes="LAL"
19..1212
/genes="LAL"
/citation={1}
/codon_start=1
/product="lysosomal acid lipase"
/protein_id="CAA83494.1"
/db_xref="GI:4456671"
/translation="MQLQGLVFFVTGILLSRVPTGTVSAVDPEVNMNVTEINRWGY
PGEHSLVTGDDGYILSIHRIPRGKNHFGKGRPVVYVYQHLGLADSSNWTNIDNSSL
GFLADRGDFVNMGNSTWSLKHRTLSVSDQEFWAFSEDEMAKYDLPASINYLK
TGQEQIYYVGHSGQCTIGFTAFSQMPELAKKIKMFLVLPVLSLNFASGPLLQGLRLP
DPLKDMFGOKQFLPOSAMUKWLSIHVCTHVIKELCANVFELLCGFNEKNLNMRLP
VYTHCPAELLQNMELHWGOVFKYKLOAFDWSSEKNEYHYNQSPFSPNINKNRPL
TALMSGROWLADINDITILLTPKLVYHKNIPEWDHLDIFWGLDAPWKLIDEIISL
MKRYQ"
sig_peptide 19..75
/genes="LAL"
/product="lysosomal acid lipase"
mat_peptide 76..1209
/genes="LAL"
/polyA_signal 2246..2251
/polyA_site 2358
BASE COUNT 604 a 549 c 524 g 681 t
ORIGIN

alignment_scores:
Quality: 1206.00 Length: 415
Ratio: 3.611 Gaps: 3
Percent Similarity: 80.482 Percent Identity: 54.217

alignment_block:
US-09-333-159-47 x MMLYACLY
Align seg 1/1 to: MMLYACLY from: 1 to: 2358
12 SerHisArgMetGluMetTrpLeuLeuValAlaTyrMetPheG1 28
4 GCCCACTACCGGAAGATGCAACTCCAGGCGCTGGTG...TTCGTTTTCAC 50
28 nArgAsnValAsnSerValHisMetProThr.....LysAlaValA 42
51 CATGGGATCTCTGTTCTCGGTGCCACGGAACTGTATCAGCTGTGG 100
42 spProGluAlaPheMetAsnIleSerGluIleIleGlnHisGlnGlyTyr 58
101 ATCCTGAAGTCAACATGAATGTGACCGAGATATCATCGCTGGGATAT 150
59 ProCysGluGluTyrGluValAlaThrGluAspGlyTyrIleLeuSerVa 75
151 CCTGGTGAAGAACACTCGCTCTGACAGGAGATGCTATATTCTGAGTAT 200
75 LasnArgIleProArgGlyLeuValGlnProLysThrLysThrGlySerArgp 92
201 TCACCGAATCTCTGCTGGTGGGAAGAACCAATTTTGGCAAGGTCCAGAC 250
92 roValValLeuLeuGlnHisGlyLeuValGlyGlyAlaSerAsnTrpIle 108
251 CAGTTGTGTATCTTCAGCACGCGCTTGTCTGGCAGATTCTAGTAACCTGGGTC 300
109 SerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAlaGlyPh 125

```

[illegible]

```

1201 AAATACCACTGAATGCTGGACTTGAATACTCCCTCCCTCAGAGGGG 1245
seq_name: gb_pat1:A26689

seq_documentation_block:
LOCUS A26689 1378 bp DNA PAT
DEFINITION Precursor of rabbit gastric lipase coding sequence.
ACCESSION A26689
VERSION A26689.1 GI:905029
KEYWORDS
SOURCE
ORGANISM synthetic construct.
artificial sequence.
REFERENCE
AUTHORS 1 (bases 1 to 1378)
TITLE Benicourt,C., Blanchard,C. and Junien,J.L.
RECOMBINANT gastric lipase from rabbit and pharmacological
compositions
JOURNAL Patént: EP 0542629-A 9 19-MAY-1993;
INSTITUT DE RECHERCHE JOUVEINAL
FEATURES
source Location/Qualifiers
1..1378
/db_xref="taxon:32630"
/organism="synthetic construct"

BASE COUNT 397 a 287 c 276 g 418 t
ORIGIN

alignment_scores:
Quality: 1167.00 Length: 402
Ratio: 3.505 Gaps: 3
Percent Similarity: 82.836 Percent Identity: 54.478

alignment_block:
US-09-333-159-47 x A26689 ..

Align seg 1/1 to: A26689 from: 1 to: 1378

16 GluMetTrpLeuLeuValAlaIaTyMetPheGlnArgAsnValas 32
|||||
51 AAAATCGGGTGTCTTTCATGGTGGCAGCTTGTCTATCTGCA...CTTGG 97
|||||
32 nSerValHis.....MetProThrLysAlaValaspp 43
|||||
98 AACTACACATCGTCTTTTGGAAAAATCAGCACCA.....AAC 138
|||||
43 roGluAlaPheMetAsnIleSerGluIleIleGlnHisGlnGlyTrp 59
|||||
139 CTGAAGCGAATATGAATATTAGTCAGATGATTTCTCTCTGGGATACCA 188
|||||
60 CysGluGluTyrgluValAlaIaThrGluAspGlyTyrrileLeuSerValas 76
|||||
189 AGTGAANAATATCAAGTTGTGACTGAAGATGGTTATATCTCTTGAGTTAA 238
|||||
76 nArgIleProArgGlyLeuValGlnProLysLysThrGlySerArgProv 93
|||||
239 CAGAATTCCTTATGGGAAGAAAAAATTCAGGGAACAGAGGCCAGAGCCG 288
|||||
93 alValLeuLeuGlnHisGlyLeuValGlyAlaSerAsnTrpIleSer 109
|||||
289 TCGTATTTTTCAGCATGGTTTGTCTGCATCAGCTCAAACTGGATCTCC 338
|||||
110 AsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAlaGlyPheas 126
|||||
339 AACCTGCCAACACACACCTTGGCTTTATTCGTCAGATGCTGGTTATGG 388
|||||
126 pValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerArgLysHisLysT 143
|||||
389 TGTGTGGCTGGGGAACAGCAGAGGAAATACCTGCTCCAGAAAGAACTTAT 438
|||||
143 hrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTyrrAspGluMet 159
|||||
439 ACTATTACACGACTCGGTGAATCTTGGGCTTTTAGCTTTGATGAAATG 488

```

```

160 AlaArgPheAspLeuProAlaValIleAsnPheIleLeuGlnLysThrGI 176
176 yGlnGluLysIleTyrValGlyTyrSerGlnGlyThrThrMetGlyP 193
539 ACAGGAAAAGCTGCACCTATGTTGGTCATTCCAGGGCACCACCATGGTT 588
193 heileAlaPheSerThrMetProGluLeuAlaGlnLysIleLysMetTyr 209
589 TTATTGCTCTTTCTACAATCCCAAGCTGGCTGAAGAAGTCAAAACCTTT 638
210 PheAlaLeuAlaProIleAlaThrValLysHisAlaLysSerProGlyTh 226
639 TATGCTATAGCTCCAGTCCGCTGCTGAGTATACAAAAGCCTGTATAA 688
226 rLysPheLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysL 243
689 CAACTTAGTATTCTCCCAACCATGTTCAAGATTATATTGTTGACA 738
243 ySGLuPheLeuTyrGlnThrArgPheLeuArgGlnLeuValIleTyrLeu 259
739 AAATATCTACCCACACAAATTTCTTGATCAATTTCTTGCCACTCAAGTG 788
260 CysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMetLeuLeu 276
789 TCTTCCCGTGACACACTGAATGTAATTTGCAGCAATGCTTATTATCAT 838
276 uGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlaSerValTyrA 293
839 TTGTGCTTTGACAGCGCAAACTTCAACATGAGTGGCTGGTGTGTAGC 888
293 IaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTrpSer 309
889 TGTCACTAATCCGGCAGGAACCTTCAGTTCAAAACATGCTGCACGTGACC 938
310 GlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlySerGluTh 326
939 CAGGCTGTTAATCTGGAATTTTCAGCTTTAATTTGGGAAGTCCAGC 988
326 rLysAsnLeuGluLysCysAsnGlnProThrProValArgTyrArgValA 343
989 TCAGAATGATGAGCTTCAATCAGCCACACCTCCCTACTACATGTA 1038
343 rGaspMetThrValProThrAlaMetTrpThrGlyGlnAspTrpLeu 359
1039 CCGCCATGAATGTGCCAATTTGAGTGTGGAGTGGTGGCAATGACTGGTTG 1088
360 SerAsnProGluAspValLysMetLeuLeuSerGluValThrAsnLeuI 376
1089 GTTGACCCCAAGATGTTGACCTTTTGTCTCCAAAACCTTCTAACTCAT 1138
376 eTyrHisLysAsnIleProGluTrpAlaHisValAspPheIleTrpGlyL 393
1139 TTACCACAGGAGATCTTCCATATAATACACTTGGATTTATCTGGGCAA 1188
393 euAspAlaProHisArgMetTyrAsnGluIleIleHisLeuMetGlnGln 409
1189 TGAATGCTCTCAAGAAGTTTCAATGAATATTATTTCTATGATGCAAAA 1238
410 GluGlu 411
1239 GATAAA 1244
seq_name: gb_pat1.A26690
seq_documentation_block:
LOCUS A26690 1378 bp DNA PAT 05-APR-1995
DEFINITION Precursor of rabbit gastric lipase (comp.).
ACCESSION A26690
VERSION A26690.1 GI:905030
KEYWORDS
SOURCE synthetic construct.

```

```

ORGANISM synthetic construct
REFERENCE artificial sequence.
AUTHORS Benicourt,C., Blanchard,C. and Junten,J.L.
TITLE Recombinant gastric lipase from rabbit and pharmaceutical
COMPOSITIONS
JOURNAL Patent: EP 0542629-A 10 19-MAY-1993;
INSTITUT DE RECHERCHE JOUVEINAL
FEATURES
source 1.1378
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 418 a 276 c 287 g 397 t
ORIGIN
alignment_scores:
Quality: 1167.00 Length: 402
Ratio: 3.505 Gaps: 3
Percent Similarity: 82.836 Percent Identity: 54.478
alignment_block:
US-09-333-159-47 x A26690/rev ..
Align seg 1/1 to reverse of: A26690 from: 1 to: 1378
16 GluMetTrpLeuLeuIleValAlaTyrMetPheGlnArgAsnValas 32
1328 AAAATGTGGTGCTTTTCATGGTGGCAGCTTTTGTCTATCTGCA...CTTGG 1282
32 nserValHis.....MetProThrLysAlaValAsp 43
1281 AACTACACATGCTCTTTTGGAAATCAGCACCAACA.....AACC 1241
43 roGluAlaPheMetAsnIleSerGluIleIleGlnHisGlnGlyTyrPro 59
1240 CTGAAGTGAATATGAATATTAGTCAGATGATTCTCTACTGGGATACCCA 1191
60 CysGluGluTyrGluValAlaThrGluAspGlyTyrIleLeuSerValas 76
1190 AGTGAATAATATGAAGTTGTGACTGAAGATGCTTATATCTTGAGGTAA 1141
76 nArgIleProArgGlyLeuValGlnProLysThrGlySerArgProv 93
1140 CAGAAATCTTATGGGAGAAAAATTCAGGGAACAGAGCCAGAGACCCG 1091
93 alValLeuGlnHisGlyLeuValGlyValSerAsnTrpIleSer 109
1090 TCGTATTTTGCAGCATGGTTTGTGTCATCAGCTCAAACTGGATCTCC 1041
110 AsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAlaGlyPheas 126
1040 AACCTGCCACACACAGCCTGGCCTTTATCTGGCAGATGCTGGTTATGG 991
126 pValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerArgLysHisLysT 143
990 TGTGTGGTGGGGAACAGCAGAGGAATACCTGGTCCAGAGAAACTTAT 941
143 hrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTyrAspGluMet 159
940 ACTATTCCACGACTCCGTTGAATCTGGCTTTTAGCTTTGATGAATG 891
160 AlaArgPheAspLeuProAlaValIleAsnPheIleLeuGlnLysThrGI 176
890 GCTAATATGACCTTCCAGCCCAATTTGACTTCATTGTAAGAAGAACTGG 841
176 yGlnGluLysIleTyrTyrValGlyTyrSerGlnGlyThrThrMetGlyP 193
840 ACAGGAAAAGCTGCACCTATGTTGGTCAATTTCCAGGGCACCACCATTTGGTT 791
193 heileAlaPheSerThrMetProGluLeuAlaGlnLysIleLysMetTyr 209
790 TTATTGCTCTTTCTACAATAATCCCAAGCTGGCTGAAGAAGTCAAAACCTTT 741

```

```

210 PheAlaLeuAlaProIleAlaThrValIySHisAlaLysSerProGlyTh 226
740 TATCATAGCTCCAGTCGCCACTGTGAAGTATACAAAAGACCTTGTA 691
226 rLysPheLeuLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysL 243
690 CAACCTTAGTGTATCTCTCAACCATGTTCAAGATTATATTGGTGACA 641
243 ySGluPheLeuTyGlnThrArgPheLeuArgGlnLeuValIleTyLeu 259
640 AATATTCTTACCCACACAATTTCTTGTCAATTCATCAATTTCTGCCACTCAAGTG 591
260 CysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMetLeuLeuLe 276
590 TGTTCCTCGAGACACGAATGTAATTTGACGAATGCCTTATTTATCAT 541
276 uGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlaSerValTyRA 293
540 TTGTGGCTTTGACAGCCCAACITTTGAACATGAGTGGCTTGGATGTGTACG 491
293 laAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTrpSer 309
490 TGTACATATTCGGCAGGACACTTCAGTTCAAAACATGCTGCACCTGGACC 441
310 GlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlySerGluTh 326
440 CAGCTGTAAATCTGGAATTTTCAAGCTTTTAATTTGGGAAGTCCAGC 391
326 rLysAsnLeuCluLysCysAsnGlnProThrProValArgTyArgVala 343
390 TCAGATGTAGTGCACTTCAATCAGCCACACCTTCCTACTACAAATGTGA 341
343 rGAspMetThrValProThrAlaMetTrpThrGlyGlyGlnAspTrpLeu 359
340 CCGCATGAATGTGCCAATTTGCAAGTGTGGAGTGGTGGCAATGACTGGTG 291
360 SerAsnProGluAspValLysMetLeuLeuSerGluValThrAsnLeuIl 376
290 GCTGACCCCAAGATGTTGACCTTTTCTTCCAAAACCTTTCTAATCTCAT 241
376 eTyHHisLysAsnIleProGluTrpAlaHisValAspPheIleTrpGlyL 393
240 TTACCAAGAGAGATTTCTCATATATCACTTGGATTATTCATCGGCCAA 191
393 euAspAlaProHisArgMetTyAsnGluIleHisLeuMetGlnGln 409
190 TGAATGCTCTCAAGAAGTTTACAATCAATTAATTTCTATGATGGCAAAA 141
410 GluGlu 411
140 GATAAA 135

```

seq_name: gb_pat1:A01157

```

seq_documentation_block: 1336 bp mRNA PAT 29-JAN-1993
LOCUS A01157
DEFINITION R.norvegicus mRNA for prelingual lipase protein.
ACCESSION A01157
VERSION A01157.1 GI:14748
KEYWORDS lipase.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1336)
AUTHORS Williamson, R.
JOURNAL Patent: WO 8500381-A 3 31-JAN-1985;
Celltech Ltd
FEATURES
Location/Qualifiers
1..1336
/organism="Rattus norvegicus"

```

CDS

```

/db_xref="taxon:10116"
7..1194
/codon_start=1
/product="prelingual lipase"
/protein_id="CAA00136.1"
/db_xref="GI:14749"
/db_xref="SWISS-PROT:P04634"
/translation="MWLLITSVITFGGAHGLFGKLGPNPEANNNISOMITYWGY
COEYEVVTDGTYLGVYRIPHGKNSINENIGKRPVYVYLOHGLIASATNWTNANLPNNSLA
FMLADAGYDVLGNSRGNTNYSKNTVYSPDSVEFWAFSDENAKYDLPATINFIQVKT
GOEKIHYVHSGOITGTIAFSTNPTLAKKIETVALAPVATVKYTSPLKISFIPT
FLFKLMFGKKMPLPHYTFDDFLGTGECVSEVLDLCSNTLFIKFCGFDKKNLNVSRDV
YLGNPAGTSVDFLHWAOQLVRSGRFOAFNMGSPONMLHYNOKTTPPEYDVSAAMTVPV
AVYNGNDILADPDQVAMLLPKLNLPHKEILAYNHLDFIWMADAPQEVYNEIMISM
AED"
BASE COUNT 409 a 283 c 265 g 379 t
ORIGIN

alignment_scores:
Quality: 1166.50 Length: 399
Ratio: 3.556 Gaps: 1
Percent Similarity: 82.206 Percent Identity: 54.637

alignment_block:
US-09-333-159-47 x A01157
Align seg 1/1 to: A01157 from: 1 to: 1336

16 GluMetTrpLeuLeuIleLeuValAlaTyMetPheGln...ArgAsnVa 31
:::|||||:::|||||:::|||||:::
4 AAGATGGCGCTGTATTATAACAAGTGTGATATCAACATTCGGAGGTCG 53
:::|||||:::|||||:::|||||:::
31 lAsnSerValHisMetProThrLysAlaValAspProGluAlaPheMetA 48
:::|||||:::|||||:::|||||:::
54 ACATGGCTATTGGAAAACCTGGGTCTCGAAACCCCTGAAGCAATATGA 103
:::|||||:::|||||:::|||||:::
48 snLseSerGluIleIleGlnHisGlnGlyTyProCysGluGluTyGlu 64
:::|||||:::|||||:::|||||:::
104 ATATTAGTCAGATGATAACTTACTGGGGATATCCATGTCAGAATATGAA 153
:::|||||:::|||||:::|||||:::
65 ValAlaThrGluAspGlyTyIleLeuSerValAsnArgIleProArgGl 81
:::|||||:::|||||:::|||||:::
154 GTTGTACTGAAGATGGCTACATCTCTGGGGGTCTACAGAATTCCTCATGG 203
:::|||||:::|||||:::|||||:::
81 yLeuValGlnProLysLysThrGlySerArgProValValLeuLeuGlnH 98
:::|||||:::|||||:::|||||:::
204 GAAGATAATCTGAAAATATAGGCAAGACACCTGTGGTGTATTGGCAGC 253
:::|||||:::|||||:::|||||:::
98 lSGlyLeuValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAsn 114
:::|||||:::|||||:::|||||:::
254 ATGTTTGATCGCATCAGCCACAACCTGGATTGCAAACTACCAACAACAC 303
:::|||||:::|||||:::|||||:::
115 SerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyAs 131
:::|||||:::|||||:::|||||:::
304 AGCTTGSCCTTTATGTAGCAGACGCTGGCTATGATGTGTGGTGGGAA 353
:::|||||:::|||||:::|||||:::
131 nSerArgGlyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAspG 148
:::|||||:::|||||:::|||||:::
354 CAGTCGGAGGAATACATGCTCCAGGAAAAATGTATATCTACTCACCAGACT 403
:::|||||:::|||||:::|||||:::
148 lNAspGluPheTrpAlaPheSerTyArgGluMetAlaArgPheAspLeu 164
:::|||||:::|||||:::|||||:::
404 CAGTTGAATTTCTGGGCTTTTCAGCTTTTCAAAATGGCTAAATATGACCTT 453
:::|||||:::|||||:::|||||:::
165 ProAlaValIleAsnPheIleLeuGlnLysThrGlyGlnGluLysIleTy 181
:::|||||:::|||||:::|||||:::
454 CCGGCCACAATAAACTTCATTGTACAGAAAACCTGGCAAGAGAAGATACA 503
:::|||||:::|||||:::|||||:::
181 rTyValGlyTyTrpSerGlnGlyThrThrMetGlyPheIleAlaPheSerT 198
:::|||||:::|||||:::|||||:::
504 CTATGTGGTCATCTCAGGGCCACCACCTATTTGTTTCATTCGCTTTCTA 553
:::|||||:::|||||:::|||||:::

```

```

198 hMetProGluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaPro 214
   || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
554 CCATCCTACACTGGCCAAATAATACAGACGTTTATGCATAGCTCCA 603
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 ILeAlaThrValLysHisAlaLysSerProGlyThrLysPheLeuLeu 231
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
604 GTTGCTACCTGGAAGTATACACAAAGTCCCTGAAAAAGATTTCATTAT 653
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
231 uProAspMetMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyrG 248
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
654 TCCTACATTCTTTTCAAGCTTATGTTGGCAAGAAATGTTCTGCCCC 703
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
248 lNThrArgPheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIle 264
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
704 ACACCTACTTGTGATGACTTCTTGGTACCAAGTGTCTCACGGGAGGT 753
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265 LeuAspGlnIleCysSerAsnIleMetLeuLeuLeuGlyGlyPheAsnTh 281
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
754 CTAGATCTTCTCTGCACCAACACTTTATCATCTCTGTGGATTTCACAA 803
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
281 rAsnAsnMetAsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuA 298
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
804 GAAAAACTTAATGTAGTGGTCTTGTATGTATCTAGGCATAATCCAG 853
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
298 laGlyThrSerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSer 314
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
854 CAGGGACATCTGTTCAGACATTTCTCCACTGGGCACACGTTGTAGATCT 903
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
315 GlyLeuLeuArgAlaPheAspTrpGlySerGluThrLysAsnLeuGlu 331
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
904 GGGAAATTTCAAGCCTCACTGGGAGGCCATCCAGAACATGTTACA 953
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
331 sCysAsnGlnProThrProValArgTyrArgValArgAspMetThrValp 348
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
954 CTACACCCAGAAAACGCTCTCTGAATATGATGTGTACGCCATGACTGTGC 1003
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 roThrAlaMetTyrThrGlyGlnAspTrpLeuSerAsnProGluAsp 364
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1004 CAGTTGAGTGTGGACGGTGGCAATGACATCTCGCTGATCCCAAGAT 1053
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
365 ValLysMetLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsnI 381
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1054 GTGCCATGCTGCTTCCCAAACTCTCCAACTCTCTGTCCATGAAGGAT 1103
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
381 eProGluThrAlaHisValAspPheIleTrpGlyLeuAspAlaProHis 398
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1104 TCTTGCTACAAATCACTGACTCATCTGGCAATGGATGGCCCTCAAG 1153
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
398 rGMetTyrAsnGluIleIleHisLeuMetGlnGlnGluThrAsn 413
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1154 AGGTTTAAATGAGATGATTCTCCATGATGCGAAGACTAAAGAAT 1200
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq_name: gb_ro2:RNLIP

```

seq_documentation_block:
LOCUS      RNLIP             1355 bp      mRNA           ROD      30-MAR-1995
DEFINITION Rat mRNA for lingual lipase.
ACCESSION X02309
VERSION   X02309.1 GI:56595
KEYWORDS  glycoprotein; lipase.
SOURCE    Norway rat.
ORGANISM  Rattus norvegicus
           Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus.
REFERENCE  1 (bases 1 to 1355)
AUTHORS   Docherty,A.J., Bodmer,M.W., Angal,S., Verger,R., Riviere,C.,
           Lowe,P.A., Lyons,A., Emtage,J.S. and Harris,T.J.
TITLE     Molecular cloning and nucleotide sequence of rat lingual lipase
           cDNA
JOURNAL   Nucleic Acids Res. 13 (6), 1891-1903 (1985)
MEDLINE   85215587
COMMENT   Data kindly reviewed (12-FEB-1986) by A.J.P. Docherty.

```

```

FEATURES             Location/Qualifiers
     source            1..1355
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     sig_peptide       49..102
                     /product="put. signal peptide (aa -18 to -1)"
     CDS               49..1236
                     /note="lipase precursor"
                     /codon_start=1
                     /protein_id="CAA26179.1"
                     /db_xref="GI:56596"
                     /translation="MWLLITSVSTFGGAGHGLFGKLGPNPEANMNISOMITTYGYP
COEYEVTEDCYILGVYRIPHGKKNSENIGKRPVYVLOHGLIASATNMIANLPNNSLA
FMLADAGYDVLGNSRGNTWERNKYYSPDSVEFWAFSDFEMAKYDLPATINFIYOKT
GOEKIHYVHSGQGTIGTIFAEISTNPITLAKKIKITFYALAPVATVKTQSPKKISIPIT
FLKLMFGKMKMFLPHTYEDDFLGTEVCSEVLDLCSNTLFLFCGFDKKNLKNVSRFDV
YLGNHPAGTSVQDFLHWAQLVRSKGFQAFNMGSPSONMLHYNOKTTPPEYDVSAMTVPV
AVWNGNDILADPDQVAMLPKLSNLLFKELIAYNHLDFIWADAPQEVYNEMISM
AED"
     mat_peptide       103..1233
                     /product="lipase (aa 1-377)"
     misc_feature       505..507
                     /note="aa 153 pot. lipid binding site"
     misc_feature       1355
                     /note="polyA site"
     BASE COUNT        400 a 291 c 276 g 387 t 1 others
     ORIGIN
alignment_scores:
  Quality: 1166.50      Length: 399
  Ratio: 3.356         Gaps: 1
  Percent Similarity: 82.206   Percent Identity: 54.637
alignment_block:
US-09-333-159-47 x RNLIP
Align seg 1/1 to: RNLIP from: 1 to: 1355

```

```

16 GluMetTrpLeuLeuLeuValAlaTyrMetPheGln...ArgAsnVa 31
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
46 AAGATGTGGCTGTATTATAACAAGTGTGATATCAACATTCGAGAGTGC 95
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
31 lAsnSerValHisMetProThrLysAlaValAspProGluAlaPheMetA 48
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
96 ACATGCGCTATTTGGAAAACCTGGTCTCGGAAACCTGAAGCAATATGA 145
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
48 snIleSerGluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGlu 64
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
146 ATATTAGTCAGATGATACTTACTGGGGATATCCATGTCAAGATATGAA 195
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
65 ValAlaThrGluAspGlyTyrIleLeuSerValAsnArgIleProArgG 81
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 GTTGTTACTCAAGATGCTACATTCCTGGGGTCTACAGAAATTCCTCATGG 245
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
81 yLeuValGlnProLysLysThrGlySerArgProValValLeuLeuGlnH 98
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
246 GAAGAATAATCTGAAAATATAGCAAGAGACCTGTGTGTATTGTTCAGC 295
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 isGlyLeuValGlyValAserAsnTrpIleSerAsnLeuProAsnAsn 114
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
296 ATGTTTTGATGGCATAGCCACAACTGGATGCAAAATCTACCAACAAC 345
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
115 SerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyAs 131
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
346 AGCTGCGCTTATGTAGCAGACGCTGGCTATGATGTGTGCTGGGAA 395
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131 nSerArgGlyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAspG 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
396 CAGTCGAGGAAATACATGGTCCAGGAAAATGTATACTACTACTCACCAGACT 445
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 lNAspGluPheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeu 164

```

```

.. |||||
446 CAGTTGAATTCGGCTTTCAGCTTTGATGAATGGCTAAATATGACCT 495
165 ProAlaValIleAsnPhelLeuGlnLysThrGlyGlnGluLysIleTy 181
496 CCCGCCAATAAATCTTATGTCACAAAACACTGGACAGAGAATACA 545
181 rTyrvAlGlyTyrsrGlnGlyThrMetGlyPheIleAlaPheSerT 198
546 CTATGTTGGTCATCTCAGGCCACCACTATTGGTTTCATTGCCCTTCTA 595
198 hrMetProGluLeuAlaGlnLysIleLysMetTyrrPheAlaLeuAlaPro 214
596 CCAATCTTACATCGCCCAAAAATCAAGAGCTTTTATGCTATTAGCTCCA 645
215 IleAlaThrValLysHisAlaLysSerProGlyThrLysPheLeuLeu 231
646 GTTGCTACCGTGAAGTATACAAAGTCCCTTGAAAAGATTTCAATTAT 695
231 uProAspMetMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyG 248
696 TCCTACATTTCTTCAAGCTTATGTTGGCAAGAAAATGTTCTGCCCC 745
248 lnThrArgPheLeuArgGlnLeuValIleTyrrLeuCysGlyGlnValle 264
746 ACACCTACTTTGATGACTTTCTCGGTACCGAAGTGTGCTCACGGAGGTT 795
265 LeuAspGlnIleCysSerAsnIleMetLeuLeuGlyGlyPheAsnTh 281
796 CTAGATCTTCTCGCACCAACACTTTATTCATCTCTGTGGATTGACAA 845
281 rAsnAsnMetAsnMetSerArgAlaSerValTyrrAlaAlaHisThrLeuA 298
846 GAAAACCTAAATGTAGTCTGTTTGTATGTATCTAGGGCATAATCCAG 895
298 laGlyThrSerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSer 314
896 CAGGACATCTGTCAGACATTTCTCCACTGGGCACACAGCTGTTAGATCT 945
315 GlyGluLeuArgAlaPheAspTrpGlySerGluThrLysAsnLeuGluLy 331
946 GGGAAATTTCAAGCTTCAACTGGGAAGCCATCCCAAGACATGTTACA 995
331 scYsAsnGlnProThrProValArgTyrrArgValArgAspMetThrValp 348
996 CTACAACCAAGAAAACGCTCTCTGAATATGATGTCTAGCCCATGACTGTC 1045
348 roThrAlaMetTrpThrGlyGlnAspTrpLeuSerAsnProGluAsp 364
1046 CAGTTGAGTGGAGCGGTGGCAATGACATCTGCTGGCTGATCCCAAGAT 1095
365 ValLysMetLeuLeuSerGluValThrAsnLeuIleTyrrHisLysAsnIl 381
1096 CTCGCCATGCTGCTCCCAAACTCTCCAACCTCTCTGTTCCATAAGGAGAT 1145
381 eProGluTrpAlaHisValaspPheIleTrpGlyLeuAspAlaProHisA 398
1146 TCTTGCTTACAAATCACTCGGACTTCATCTGGGCAATGGATGCCCTCAAG 1195
398 rGmTyrrAsnGluIleIleHisLeuMetGlnGlnGluGluThrAsn 413
1196 AGGTTTACAAATGAGATGATTTCCATGATGGCAGAGAAGATAAAGAAAT 1242
seq_name: gb_om:CFLIPASE
seq_documentation_block:
LOCUS CFLIPASE 1651 bp mRNA MAM 18-JUN-1998
DEFINITION Canis familiaris mRNA for lipase.
ACCESSION Y13899
VERSION Y13899.1 GI:2204112
KEYWORDS DGL gene; gastric lipase; triacylglycerol lipase.
SOURCE dog.
ORGANISM Canis familiaris

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Vagenay,S., Joliff,G., Bertaux,O., Toselli,E., Devignes,M.D. and
Benicourt,C.
The complete cDNA sequence encoding dog gastric lipase
DNA Seq. 8 (4), 257-262 (1998)
99450174
REFERENCE 2 (bases 1 to 1651)
AUTHORS Benicourt,C.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-1997) C. Benicourt, Ecole Normale Supérieure de
Cachan, L.I.R.B.A (Laboratoire Interdisciplinaire de Recherche en
Biologie Appliquée), 61 avenue du Président Wilson, F-94235 Cachan
Cedex, FRANCE
FEATURES
Location/Qualifiers
1..1651
/organism="Canis familiaris"
/db_xref="taxon:9615"
/dev_stage="adult"
/tissue_type="stomach mucosa"
54..1250
/gene="DGL"
54..1250
/gene="DGL"
/EC_number="3.1.1.3"
/codon_start=1
/product="triacylglycerol lipase"
/protein_id="CAA74198.1"
/db_xref="GI:2204113"
/db_xref="SWISS-PROT:P80035"
/translation="MWLLTAASVISTLTGTHGLFGLKHLPTNPEVTMNIQSOMITYWGY
PAEYEVVTEDGYILGIDRIPYGRKNSENIGRPVAFLOHGLLASATNWSINLNNLSL
AFILADAGDVWLGNSRGNTWARRNLYYSPDSVEFWAFSDEMAYDLDPATIDFILKK
TGODKLHVGHSGQTTIGTAFSTNPKAKRIKTEVALAPVATVKYTTETLLAKMLVP
SFLKLIIGNKIFYPHHPFDQPLATEVCSRETYDLNCSNALFIICGFDTMNLMSRLD
VILSHNPAGTSQNVNLHWSQAVKRFQAFDMGSPVNNMHHTHQSPPYINLTDHVP
IAVWNGNDLLADPHDVLDDLKLPNLIYHRKIPYPYNNHLDIFWAMDPAQVAYNEIVSM
MGTDNK"
54..110
/sig_peptide
54..110
/gene="DGL"
/evidence="experimental"
111..1247
/mat_peptide
1602..1607
/polya_signal
432 a 419 c 352 g 448 t
BASE COUNT
ORIGIN
alignment_scores:
Quality: 1154.00 Length: 398
Ratio: 3.486 Gaps: 5
Percent Similarity: 83.166 Percent Identity: 54.271
alignment_block:
US-09-333-159-47 x CFLIPASE ..
Align seg 1/1 to: CFLIPASE from: 1 to: 1651
17 MetTrpLeuLeuIleLeuValAlaTyrrMetPheGlnArgAsnValAsnSe 33
|||||
54 ATGTGGCTGCTATTAAACAGCGCGGAGTGTGATATCT...ACACTTGGAAAC 100
33 rValHis.....MetProThrLysAlaValAspProG 44
|||||
101 CACACATGGTTATTTCGAAAATTTACATCCACA.....AACCTG 141
44 luAlaPheMetAsnIleSerGluIleIleGlnHisGlnGlyTyrrProCys 60
|||||
142 AAGTGACCATGAATATAAGTCAGATGATCAGCTACTGGGGATACCCAGCT 191
61 GluGluTyrrGluValAlaThrGluAspGlyTyrrIleLeuSerValAsnAr 77
|||||

```

```

376 eTyrHslyAsnIleProGluTrpAlaHisValAspPheiletrpGlyL 393
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1139 TTACCACAGGAAGATTCCCTTCATTACAATCACTTGGACTTATCTGGGCCA 1188

393 euAspAlaProHisArgMetTyrAsnGluIleHisLeuMet 407
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1189 TGGATGCCCTCAAGCGGTTTACAATGAATTCCTTCCCATGATG 1232

seq_name: gb_pat1:AR039022

seq_documentation_block:
LOCUS AR039022 1137 bp DNA PAT 29
DEFINITION Sequence 2 from patent US 5807726.
ACCESSION AR039022
VERSION AR039022.1 GI:5958385
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1. (bases 1 to 1137)
AUTHORS Blanchard, C., Benicourt, C. and Junien, J.
TITLE Nucleic acids encoding dog gastric lipase and their
production of polypeptides
JOURNAL Patent: US 5807726-A 2 15-SEP-1998;
FEATURES Location/Qualifiers
source 1..1137
/organism="unknown"
BASE COUNT 296 a 308 c 249 g 284 t
ORIGIN

```

```
alignment_scores:
  Quality: 1141.50      Length: 372
  Ratio: 3.624          Gaps: 3
  Percent Similarity: 84.677  Percent Identity: 56.452

alignment_block:
  US-09-333-159-47 x AR039022 ..

Align seg 1/1 to: AR039022 from: 1 to: 1137
```

170 heileLeuGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSer 186
 410 TCATCTTGAAGAAACGGGACAGACAGTACACTACGTTGGCCATTC 459
 187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
 460 CAGGCGACCAACCATGTGTTTCATCGCCTTTCCACCAATCCCAAGCGGC 509
 203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
 510 GAAACGGATCAAAACCTTCTATGATAGCTCCGTTGCCACCGTGAAT 559
 220 IsAlaLysSerProGlyThrLysPheLeuLeuProAspMetMetile 236
 560 ACACCGAAACCTGTTAAACAAACTCATGCTCGTCCCTGCTCTCTTC 609
 237 LysGlyLeuPheGlyLysGluPheLeuTyrGlnThrArgPheLeuAr 253
 610 AAGCTTATATTGGAACAAATATTC...TACCACACCACTCTTTGA 656
 253 gGln...LeuValIleTyrLysGlyGlnValIleLeuAspGlnIleC 269
 657 TCAATTTCTGCCACCGAGGTATGCTCCCGGAGCGGTGATCTCTCT 705
 269 ysSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAsnMetAsn 285
 707 GCACAAACGCCCTGTTTATCATTTGTGGATTGACACTATGAAC 756
 286 MetSerArgAlaSerValTyrAlaIleHisThrLeuAlaGlyThrSera 302
 757 ATGAGTCGCTGGATGTATCTGTCTACATTAATCCACGACGACATCGGT 806
 302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgA 319
 807 TCAGACGTGCTCCACGTGCTCCAGGCTGTTAAGTCTGGGAAGTCCAAG 856
 319 laPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
 857 CTTTGTACTGGGAAGCCCACTGCTGAGACATGATGCACTATCATCAGAGC 906
 336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
 907 ATGCTCTCCCTACTACACCTGACAGACATGCTGTCATTCGACGHTG 956
 352 pThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
 957 GAACGTTGGCAACGACTGTCTGGCCGACCTCAGCATGTTGACCTTTTC 1006
 369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
 1007 TTTCCAGCTCCCAATCTCTATTACACAGGAAGATTCTCTCTTACAAT 1056
 386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnG 402
 1057 CACTGTGACTTATCTGGCGCATGATGCTCCCTCAACCGGTTTACAATGA 1106
 402 uilleIleHisLeuMet 407
 1107 AATTGTTTCCATGATG 1122

seq_name: gb_pat1:AR092633

seq_documentation_block:

LOCUS AR092633 1137 bp DNA

DEFINITION Sequence 2 from patent US 5998189.

ACCESSION AR092633

VERSION AR092633.1 GI:10019385

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1137)

08-SEP-2000

AUTHORS Blanchard,C., Benicourt,C. and Junien,J.
 TITLE Polypeptide derivatives of dog gastric lipase and pharmaceutical
 compositions containing same
 JOURNAL Patent: US 5998189-A 2 07-DEC-1999;
 FEATURES Location/Qualifiers
 source 1..1137
 BASE COUNT 296 a 308 c 249 g 284 t
 ORIGIN

alignment_scores:
 Quality: 1141.50 Length: 372
 Ratio: 3.624 Gaps: 3
 Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:

US-09-333-159-47 x AR092633 ..

Align seg 1/1 to: AR092633 from: 1 to: 1137

37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleI 53
 19 CCACACA.....AACCTGAAGTGACCATGAATATAAGTCAGATGAT 59
 53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaLalThrGluAspG 70
 60 CACCTACTGGGATACCCAGCTGAGGAATATGAAGTTGTGACCGAAGACG 109
 70 LysTrpIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
 110 GTTATATCTTGGGATCGACAGAAATTCCTTATGGGAGGAAATAATTCAGAG 159
 87 LysThrGlySerArgProValValLeuGlnHisGlyLeuValGlyL 103
 160 AATATAGCGCGGAGACCTGTGTGATTTTGCACACACGCTTGTCTCGCATC 209
 103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
 210 AGCCACAACTGGATCTCCAACTGCGCCAAACACAGCTGGCTTCATCC 259
 120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
 260 TGCCCGACGCGCGGTAGCAGCTGTGGTGGGAAACACAGCGGGCAACACC 309
 137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
 310 TGGCCGACGAGGAGATCTGTACTACTCGCCGACTCCGTCGAATTCGGGC 359
 153 aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnP 170
 360 TTTTCAGCTTTGACGAGATGGCTAAATATGACTTCCGCCACCACTTGCAT 409
 170 heileLeuGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSer 186
 410 TCATCTTGAAGAAACGGGACAGACAGTACACTACGTTGGCCATTC 459
 187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
 460 CAGGCGACCAACCATGTGTTTCATCGCCTTTCCACCAATCCCAAGCTGGC 509
 203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
 510 GAAACGGATCAAAACCTTCTATGATAGCTCCGTTGCCACCGTGAAT 559
 220 IsAlaLysSerProGlyThrLysPheLeuLeuProAspMetMetile 236
 560 ACACCGAAACCTGTTAAACAAACTCATGCTCGTCCCTGCTCTCTTC 609
 237 LysGlyLeuPheGlyLysGluPheLeuTyrGlnThrArgPheLeuAr 253
 610 AAGCTTATATTGGAACAAATATTC...TACCACACCACTCTCTTTGA 656

263 TGGCCGACGCGGATGACAGCTGTGGTGGGGAACACAGGGGACACACC 312
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTpaL 153
313 TGGCCGACGAGGAATCTGTACTACTCGCCGACATCCGTCGAATTCCTGGC 362
153 aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnP 170
363 TTTCCAGCTTTCCACGAGATGGCTAAATATGACCTTCCCGCCACCATGACT 412
170 heIleLeuGlnLysThrClyGlnGluLysIleTyrTyrValGlyTyrSer 186
413 TCATCTTGAAGAAACAGGACAGGACAGCAAGCTACACTACGTTGGCCATTC 462
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
463 CAGGCGACCAACATGGTTTCATCGCTTTTCCACCAATCCCAAGCTGGC 512
203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
513 GAAACGATCAAAACCTCTATGCAATAGCTCCCGTTGCCACCGTGAAT 562
220 IsAlaLysSerProGlyThrLysPheLeuLeuLeuLeuProAspMetMetIle 236
563 ACACCGAACCCTGTTAAACAACATCATGCTCGTCCCTTCGTTCCCTCTTC 612
237 LysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArgPheLeuAr 253
613 AAGCTTATATTGGAACAAAATATTC...TACCCACACCACTTCCTTGA 659
253 gGln...LeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
660 TCAATTTCGCCACCGAGGTATGCTCCCGGACAGCGGTGGATCTCCTCT 709
269 ysSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAspMetAsn 285
710 GCACGACGCCCTGTTATCATTTGTGGATTGACACTATGAACCTTGAC 759
286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVa 302
760 ATGAGTCGCTGGATGTATCTCTCACAATAATCCACGAGGAACATCGGT 809
302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgA 319
810 TCAGAAGCTGCTCCACCTGGTCCCGAGGTGTAAAGTCTGGGAAGTTCCAAG 859
319 laPheAspTrpGlySerClyThrLysAsnLeuGluLysCysAsnGlnPro 335
860 CTTTGTGACTGGGAAGCCCAAGTTCAGAACATGATGTCACTATCATCACAGC 909
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
910 ATGCCTCCCTACTACAACCTGACAGACATGCATGTGCCAATCGCAGTGTG 959
352 pThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
960 GAACGGTGGCAACGACTGTGTCGCCGACCTTCACGATGTTGACCTTTTGC 1009
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
1010 TTTCCAGCTCCCAATCTCATTTACCACAGGAAGATTCCTCTCTACAAT 1059
386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGl 402
1060 CACTTGGACTTATCTGGGCCATGATGCCCTCAACGCGGTTTACATGA 1109
402 uIleIleHisLeuMet 407
1110 AATTGTTTCCATGATG 1125

OM of: US-09-333-159-47 to: N_Geneseq_0601.* out_format : pfs
 Date: Sep 1, 2001 9:32 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlpl
 -Q/cg2_1/USPTO_spool/US09333159/runat_31082001_122331_1419/app_query.fasta_1.488
 -DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=p2n.rng -GAPOP=12.000
 -GAPEX=4.000 -MINMATCH=0.050 -LOOPEX=0.000 -LOOPEXT=0.000
 -GAPOP=4.000 -GAPEX=0.050 -GAPOP=10.000 -GAPEXT=0.500
 -GAPOP=6.000 -GAPEX=7.000 -GAPOP=10.000 -GAPEXT=0.500
 -DELOP=6.000 -DELEX=7.000 -START=1 -MATRIX=blotsum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM_ext -MINLEN=0 -MAXLEN=200000000
 -USER=US09333159_ECGN1_1_209 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:

Query: US-09-333-159-47
 Query length: 423
 Database: N_Geneseq_0601.*
 Database sequences: 730101
 Database length: 313950809
 Search time (sec): 113.380000

score list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAF45132 +		2247.00	4488.65	4.8e-242	1269
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAF45131 +		2247.00	4483.08	9.7e-242	2044
/SID88/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ42310 +		1167.00	2315.42	5.3e-121	1378
/SID88/gcgdata/geneseq/geneseq/NA1985.DAT:AAQ50385 +		1158.50	2298.69	4.5e-120	1336
/SID88/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ68388 +		1141.50	2262.93	4.5e-118	1528
/SID88/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ58915 +		1141.50	2262.93	4.5e-118	1528
/SID88/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ68389 +		1141.50	2262.93	4.5e-118	1531
/SID88/gcgdata/geneseq/geneseq/NA1986.DAT:AAQ60685 +		1124.00	2229.03	3.5e-116	1367
/SID88/gcgdata/geneseq/geneseq/NA1986.DAT:AAQ60686 +		1124.00	2229.03	3.5e-116	1367
/SID88/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ58916 +		1124.00	2229.03	3.5e-116	1367
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ28689 +		695.00	1366.67	3.7e-68	1308
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ28679 +		691.00	1365.95	4.1e-68	699
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ44349 +		596.00	1176.33	1.9e-57	617
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ44348 +		514.50	1000.44	9.4e-48	1718
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ44347 +		502.00	977.01	7.1e-44	1483
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ44346 +		473.50	930.82	7.1e-44	572
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ44345 +		465.00	902.95	2.5e-42	1438
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ44344 +		398.00	774.21	3.7e-35	859
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ44343 +		309.50	598.52	2.3e-25	705
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ60606 +		271.50	473.86	2.0e-18	43676
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ38827 +		230.50	442.55	1.1e-16	549
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ12152 +		173.50	324.94	4.0e-10	707
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ34954 +		160.50	302.63	6.9e-09	509
/SID88/gcgdata/geneseq/geneseq/NA1998.DAT:AAQ34341 +		155.00	302.85	5.9e-09	178
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ34950 +		150.00	276.97	1.9e-07	751
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ34955 +		138.00	264.66	9.0e-07	273
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ34951 +		138.00	254.58	3.3e-06	647
/SID88/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ28743 +		117.00	208.13	0.0013	927
/SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ73350 +		117.00	198.98	0.0041	2028
/SID88/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ28741 +		116.50	207.99	0.0013	861
/SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ73347 +		116.50	207.99	0.0013	861
/SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ73384 +		110.00	173.36	0.1099	5445
/SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ73289 +		105.00	177.07	0.0682	1676
/SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ62362 +		105.00	177.07	0.0682	1676
/SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ73322 +		102.00	176.61	0.0724	1041
/SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ28252 +		99.50	131.52	23.50	32042
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ30011 +		99.50	131.52	23.50	32042
/SID88/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ89352 +		98.50	168.23	0.2120	1167
/SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ27118 +		98.00	164.97	0.3224	1416
/SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ27117 +		98.00	155.23	1.12	3257

/SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ81263 + 98.00 153.51 1.40
 /SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ69112 + 96.50 147.50 2.88
 /SID88/gcgdata/geneseq/geneseq/NA1998.DAT:AAQ35458 + 95.00 156.98 0.8983
 /SID88/gcgdata/geneseq/geneseq/NA1998.DAT:AAQ35473 + 95.00 152.47 1.60
 /SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ08414 + 95.00 152.19 1.66

seq_name: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAF45132

seq_documentation_block:

ID: AAF45132 standard; cDNA; 1269 BP.

XX AAF45132;

DT 30-MAR-2001 (first entry)

DE Human TANGO 294 ORF.

XX Gene therapy: TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hyperextension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety; ss.

OS Homo sapiens.

XX W0200077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNTUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI: 2001-032313/04.

XX P-PSDB: AAB66065.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease

XX Claim 1; Fig 6; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome, disorders,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.

XX Sequence :269 BP; 358 A; 268 C; 309 G; 334 T; 0 other;

alignment_scores:

Quality: 2247.00 Length: 423
 Ratio: 5.312 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-333-159-47 x AAF45132

Align seg 1/1 to: AAF45132 from: 1 to: 1269

```
1 MetLeuGluThrLeuSerArgGlnTrpIleValSerHisArgMetGluMe 17
|||||
1 ATGTTGGAACCTTGTCAAGACAGTGGATTGTCTCACACAGAATGGAAAT 50
|||||
17 tTrpLeuLeuIleLeuValAlaIleTyrMetPheGlnArgAsnValAsnSerV 34
|||||
51 GTGGCTTCTGATTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAG 100
|||||
34 alHisMetProThrLysAlaValAspProGluAlaPheMetAsnIleSer 50
|||||
101 TACATATGCCAACTTAAGCTGTGACCCAGACAGCATTCATGAATATTAGT 150
|||||
51 GluIleLeuGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaIle 67
|||||
151 GAAATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAAC 200
|||||
67 rGluAspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuValG 84
|||||
201 TGAAGATGGGTATATCTCTTGTAAACAGGATTCTCGAGGCCCTAGTGC 250
|||||
84 InProLysLysThrGlySerArgProValValLeuLeuGlnHisGlyLeu 100
|||||
251 AACCTAGAAGACAGGTTCCAGGCCCTGTGGTTACTGCACATGGCCCTA 300
|||||
101 ValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuG1 117
|||||
301 GTTGGAGGTGTAGCAACTGGATTTCCAACCTGCCCAACAATAGCCTGGG 350
|||||
117 yPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgG 134
|||||
351 CTTTCATTCTGGCAGATCGTGGTTTACGCTGTGGATGGGGAACAGCAGGG 400
|||||
134 lyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAspGlnAspGlu 150
|||||
401 GAAACGCTGTCTCGAAACACACAGACACTCTCCATAGACCAAGATGAG 450
|||||
151 PheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeuProAlaIa 167
|||||
451 TTCCTGGGCTTTCAGTTATGATGAGATGGCTAGGTTTGACCTTCCTGCAGT 500
|||||
167 lIleAsnPheIleLeuGlnLysThrGlyGlnGluLysIleTyrTyrValG 184
|||||
501 GATAAACTTATTTTGCAGAAACCGCCAGGAGAAAGATCTATTATGTGCG 550
|||||
184 lyTyrSerGlnGlyThrThrMetGlyPheIleAlaPheSerThrMetPro 200
|||||
551 GCTATTTCACAGGGCACCACCATGGCTTTATTGCAATTTCCACCATGCCA 600
|||||
201 GluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaIle 217
|||||
601 GAGCTGGCTCAGAAATCAAAATGATTTTGTCTTAGCACCCATAGCCAC 650
|||||
217 rValLysHisAlaLysSerProGlyThrLysPheLeuLeuLeuProAspM 234
|||||
651 TGTTAAGCATGCAAAAGCCCGGAGCAAAATTTTGTGTCGCCAGATA 700
|||||
234 etMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArg 250
|||||
701 TGATGATCAAGGGATTCTTTGGCAAAAGAAATTTCTGTATCAGACCAGA 750
|||||
251 PheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspG1 267
|||||
751 TTTCTCAGACAACCTGTATTATCTTACCTTTGTGGCCAGGTGATCTTGATCA 800
|||||
267 nIleCysSerAsnIleMetLeuLeuLeuGlyPheAsnThrAsnAsnM 284
|||||
801 GATTGTAGTATATCATGTTACTTCTGGTGGATTCACACCAACAATA 850
|||||
284 etAsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThr 300
|||||
851 TGAACATGAGCGGAGCAAGTGTATGCTGCCACACACTCTTGCTGGAACA 900
|||||
```

```
301 SerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLe 317
|||||
901 TCTGTGCAAAATATTCTACACTGGAGCCAGGCACCTGAATTCGTGTAAC 950
|||||
317 uArgAlaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnG 334
|||||
951 CCGGGCATTTTACTGGGGAGTGAGACCAAAATCTGGAAAATGCAATC 1000
|||||
334 InProThrProValArgTyrArgValArgAspMetThrValProThrAla 350
|||||
1001 AGCCAACCTCTGTAAAGGTACAGAGTCAGAGATATGACGGTCCCTACAGCA 1050
|||||
351 MetTrpThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMe 367
|||||
1051 ATGTGGACAGGAGGTCAAGACTGGCTTTCAAATCCAGAAGACGTGAAAAT 1100
|||||
367 tLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluT 384
|||||
1101 GCTGCTCTCTGAGGTGACCAACCTCATCTACCAATAAGAATATCTCTGAAT 1150
|||||
384 rPalaHisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyr 400
|||||
1151 GGGCTCACGTGGATTTCATCTGGGGTTTGGATGCTCTCACCGTATGTAC 1200
|||||
401 AsnGluIleIleHisLeuMetGlnGlnGluThrAsnLeuSerGlnG1 417
|||||
1201 AATGAATCATCATCTCATCTGATGCAGCAGGAGGAGACCAACCTTTCCCAAGG 1250
|||||
417 yArgCysGluAlaValLeu 423
|||||
1251 ACGGTGTGAGGCCGTATTG 1269
|||||
```

seq_name: /SIDSe/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF45131

seq_documentation_block:

ID AAF45131 standard; cDNA; 2044 BP.

XX AAF45131;

XX 30-MAR-2001 (first entry)

DE Human TANGO 294 cDNA.

XX Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety; ss.

OS Homo sapiens.

XX WO200077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI: 2001-032313/04.

XX P-PSDB; AAB66065.

PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
screening assays and diagnostic assays and for the treatment of
neurological diseases such as Alzheimer's, Parkinson's and Huntington's
disease.

XX Claim 1; Fig 6; 359pp; English.

xx The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF4511-145136 and AAF45137-145139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 xx

Sequence 2044 BP; 618 A; 401 C; 460 G; 565 T; 0 other;

alignment_scores:
 Quality: 2247.00 Length: 423
 Ratio: 5.312 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-333-159-47 x AAF45131 ..

Align seg 1/1 to: AAF45131 from: 1 to: 2044

1 MetLeuGluThrLeuSerArgGlnTrpIleValSerHisArgMetGluMe 17
 126 ATGTTGGAACCTTGCAAGACAGTGGATTGCTCACACAGAATGGAAT 175
 17 trpIleuLeuIleLeuValAlaTyrMetPheGlnArgAsnValAsnServ 34
 176 GTGGCTTCGATTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAG 225
 34 aHisMetProThrLysAlaValAspProGluAlaPheMetAsnIleSer 50
 226 TACATATGCCAACTAAAGCTGTGGACCCAGAACGATTCATGAATATTAGT 275
 51 GluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaLath 67
 276 GAAATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAATCGCAAC 325
 67 rGluaspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuValG 84
 326 TGAAGATGGGTATATCTCTTCTGTTAACAGGATTCCTCGAGGCTTAGTGC 375
 84 InProLysLysThrGlySerArgProValValLeuLeuGlnHisGlyLeu 100
 376 AACCTAAGAAACAGAGTTCAGGCTGTGGTGTACTGCAGCATGGCCTA 425
 101 ValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuG1 117
 426 GTTGAGGTGTGTAGCACTGGATTTCCAACTGCGCCCAACATAGCTGGG 475
 117 yPheIleLeuAlaaspAlaGlyPheaspValTrpMetGlyAsnSerArgG 134
 476 CTTCAATCTGGCAGATGCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 525
 134 lYAsnAlaTrpSerArgLysHisLysThrLeuSerIleAspGlnAspGlu 150
 526 GAAACGCTGGTCTCGAAACACAGACACTCTCCATAGACCAAGATGAG 575
 151 PheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeuProAlaVal 167
 576 TTCTGGGCTTTTCAGTTATGATGAGATGGCTAGGTTTGACCTTCCTGCAGT 625
 167 lIleAsnPheIleLeuGlnLysThrGlyGlnGluLysIleTyrTyrValG 184
 626 GATAAAGCTTTATTTGCAAGAAACGGCCAGGAGAAAGATCTATTATGTGC 675
 184 lYTyTyrSerGlnGlyThrThrMetGlyPheIleAlaPheSerThrMetPro 200
 |||||||

676 GCTATTTCACAGGGCACCACCATGGGCTTTATTGTCATTTTCCACCATGCCA 725
 201 GluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaTh 217
 |||||||
 726 GAGCTGCTCAGAAAATCAAAATGATTATTGCTTTAGCACCCATAGCCAC 775
 217 rVallyHisAlaLysSerProGlyThrLysPheLeuLeuLeuProAspM 234
 |||||||
 776 TGTAAAGCATGCAAAAACCCCGGACCAATTTTGTGTGCGCCAGATA 825
 234 etMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArg 250
 |||||||
 826 TGATCAATCAAGGATTGTTGGCAAAAAGAATTTCTGTATCAGACCAGA 875
 251 PheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspG1 267
 |||||||
 876 TTTCTCAGACAACTTGTATTATTACCTTTGTGGCCAGGTGATTCTTGATCA 925
 267 nIleCysSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAspM 284
 |||||||
 926 GATTGTAGTAATATCATGTTTACTTCTGGTGGATTCAACACCAACAATA 975
 284 etAsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThr 300
 |||||||
 976 TGAACATGAGCGAGCAAGTGTATATGCTGCCACACTCTTGTGGAACA 1025
 301 SerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLe 317
 |||||||
 1026 TCTGTGCAAAATATTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAAC 1075
 317 uArgAlaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAspG 334
 |||||||
 1076 CGGGCATTTGACTGGGGAGTGAGACCAAAAATCTGGAATAATGCAATC 1125
 334 InProThrProValArgTyrArgValArgAspMetThrValProThrAla 350
 |||||||
 1126 AGCCAACCTCCGTAAAGTACAGAGTCAGAGATATGACGGTCCCTACAGCA 1175
 351 MetTrpThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMe 367
 |||||||
 1176 ATGTGACAGAGGAGTCCAGACTGGCTTTTCAATCCAGAGAGAGTGAAT 1225
 367 tLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluT 384
 |||||||
 1226 GCTGCTCTCTGAGGTGACCAACCTCATCTACCATAGAATATTCCTGAAT 1275
 384 tpAlaHisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyr 400
 |||||||
 1276 GGGCTCAGCTGGATTTCATCTGGGGTGTGGATGCTCCTCACCGTATGTAC 1325
 401 AsnGluIleIleHisLeuMetGlnGlnGluThrAsnLeuSerGlnG1 417
 |||||||
 1326 RATGAATCATCTCATCTGATGCGAGGAGGAGAGACCAACCTTTCCAGGG 1375
 417 yArgCysGluAlaValLeu 423
 |||||||
 1376 ACGGTCTGAGGCGGTATTG 1394

seq_name: /SID8/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ42310

seq_documentation_block:

ID AAQ42310 standard; DNA; 1378 BP..

XX

AC AAQ42310;

XX

DT 20-SEP-1993 (first entry)

XX RGL precursor.

XX

XX Rabbit gastric lipase; RGL; pRGLN2.1; fat; bioconversion;

KW hydrolysis; transesterification; ds.

XX

OS Oryctolagus cuniculus.

410 GluGlu 411

• •
• •
• •
• •
• •
• •

seq_name: /SIDS8/gcgdata/geneseq/geneseq/NA1985.DAT: AAN50385

seq_documentation_block:

ID AAN50385 standard; DNA; 1336 BP.

AC AAN50385;

DT 17-JAN-1992 (first entry)

Rat prelingual lipase gene.

XX
KW
Prelinqual lipase; enzyme: EC-3.1.1.3: ss.

Rattus rattus.

XX	Key	Location/Qualifiers
EH		

EH	key	location
FT	CDS	7-1194

```

CDS
/...1134
/*tag= a

```

PN GB2142337-A.

16-JAN-1985.

XX
PF 29-JUN-1984: 84GR-0016581

XX
DD 01-III-1993. 92CB-

PR 01-JUL-1983; 83GB-001/989.
PR 05-SEP-1983; 83GB-0023759

XX
PA
(CELL:-) CELLTECH LTD

XX Carey NH. Williamson R. PT

XX
WDI: 1985-01450/03

DR WPI: I983-UI44307
DR P-PSDB: AAP50322

XX New lingual lipase protein for treatment of lipase deficiency - also
PT new pre-lingual lipase protein and related products
PT

PS Disclosure; Fig 5; 15pp; English.

CC This sequence may be expressed in a transformant host organism using
CC a vector plasmid in order to produce a prelingual lipase protein.
CC The expressed protein may be used for the treatment of lipase
CC deficiency.

Sequence 1336 BP; 409 A; 285 C; 263 G; 379 T; 0 other:

alignment scores:

Quality: 1158.50 Length: 399

Ratio: 3.543
Gaps: 1

Percent Similarity: 81.955 Percent Identity: 54.386

alignment_block:

US-09-333-159-47 x AAN50385

Align seg 1/1 to: AAN50385 from: 1 to: 1336

16 GluMetTrpLeuLeuValAlaTyrMetPheGln...ArgAspVa 31

[illegible]

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

31 lAsnSerValHisMetProThrLysAlaValAspProGluAlaPheMetA 48

54 ACATGGCCCTATTGGAAACTGGGTCTGGAAACCTTGAAGCAAAATATGA 10

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040

16-DEC-1992; 92FR-0015201.
(LJOU) INST RECH JOUVEINAL.
Benicourt C, Blanchard C, Junien J;
WPI: 1994-217890/26.
P-PSDB; AAR56871.
Recombinant canine gastric lipase and nucleic acid encoding it -
are used for improving absorption of ingested fat, treating
mucoviscidosis etc. and in enzymatic bio-conversions
Claim 2; Fig 8; 52pp; French.
The sequence given below is the sequence of figure 8, altered
according to the amendments described on page 2 of the appended
letter.
CGL is used to improve absorption of ingested fat, in healthy and
sick patients (e.g. having altered levels of gastric lipase); to
treat conditions associated with insufficiency (or lack) of lipases,
esp. mucoviscidosis or exocrine pancreatic insufficiency and partic.
where immobilised, for bioconversions, e.g. hydrolysis or
transesterification (other mammalian gastric lipases, or derivs.,
can be used in this application).
Sequence 1531 BP: 395 A; 386 C; 329 G; 421 T; 0 other;

```
alignment_block:
US-09-333-159-47 x AAQ68389 ..
```

37 ProThrLysAlaValaspProGluAlaPheMetasNileSerGluIleil 53
||||| :|||:::
22 CCCACA AACCCCTGAAGTCACCATGAATATAAGTCAGATGAT 62

53 eGlnHisGlnGLyTyr-ProCysGluGluTyrGluValAlaThrGluAspG 70
:|||||
63 CACCCTCTGGGGATACCCCACTCTGGGATATATGACTGTGCACCGAAGACG 112

FF 10-DEC-1993; 93MO-FR01200.


```

|||||
323 TGGATTTCACACCTGCCAACACAGCTTGCTTCATTCGGCAGATGC 372
123 aGlyPheAspValTTPMetGlyAsnSerArgGlyAsnAlaTrpSerArgL 140
|||||
373 TGGTTATGATGTGGCTGGGCAACAGCAGAGCAACACCTGGGCGAGAA 422
140 ySHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTyr 156
|||||
423 GAACTTGTTACTATTCCACCATTCAGTTGATTCGGCTTCAGCTTT 472
157 AspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLeuG1 173
|||||
473 GATGAATGGCTAAATATGACCTCCAGCCCAATCGACTTCATTGTAAA 522
173 nLysThrGlyGlnGluLysIleTyrValGlyTyrSerGlnGlyThrT 190
|||||
523 GAAACTGGACAGAGCAGCTACATATGTTGGCCATTCCTCCAGGGCACA 572
190 hrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLysIle 206
|||||
573 CCATTGGTTTATTGCTTTCCACCAATCCAGCCTGGCTAAAGAATC 622
207 LysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLysSe 223
|||||
623 AAAACCTTCTATGCTCTAGCTCTGCTCCACTGTGAAGTATACAAAAG 672
223 rProGlyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyLeuP 240
|||||
673 CTTTATAACAACCTAGATTGTTGCTCAATCCCTCTCAAGTTTATAT 722
240 hGlyLysLysGluPheLeuTyrGlnThrArgPheLeuArgGlnLeuVal 256
|||||
723 TTGGTGACAAATAATTACCCACACAACCTCTTTGATCAATTTCTTGT 772
257 IleTyrLeuCysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMe 273
|||||
773 ACTGAAGTGTGCTCCCGTAGATGCTGAATCTCTTTCGACCAATGCCT 822
273 tLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlas 290
|||||
823 ATTTATAATTTGGATTTGACAGTAAGACTTTAACACGAGTCCTTGG 872
290 erValTyrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeu 306
|||||
873 ATGTGTATCTATCATAATAATCCAGCAGGAACCTCTGTTCAAAACATGTC 922
307 HisTrpSerGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpG1 323
|||||
923 CAITGGACCCAGCGCTTAAAGTCTGGAAATCCCAAGCTTATGACTGGGG 972
323 ySerGluThrLysAsnLeuGluLysCysAsnGlnProThrProValArgT 340
|||||
973 AAGCCAGTTGAGATAGGATGACATATGATCAGTCCCAACCTCCCTACT 1022
340 yArgValArgAspMetThrValProThrAlaMetTrpThrGlyGln 356
|||||
1023 ACAATGTGACACCCATGAATGTACCAATTTGCAGTGTGAAGCGTGCAAG 1072
357 AspTrpLeuSerAsnProGluAspValLysMetLeuLeuSerGluValTh 373
|||||
1073 GACCTGTGGCTGACCCCAAGATGTTGGCTTTTGTCTCCAAACTCCC 1122
373 rAsnLeuIleTyrHisLysAsnIleProGluTrpAlaHisValAspPheI 390
|||||
1123 CAATCTTTTACCACAGGAGATTCCTTTTACAACTACCTTGGACTTTA 1172
390 leTrpGlyLeuAspAlaProHisArgMetTyrAsnGluIleIleHisLeu 406
|||||
1173 TCTGGCAATGGATGCCCTCAAGAAGTTTACAATGACATGTTCTTCTATG 1222
407 MetGlnGlnGlu 411
|||||

```

```

1223 ATATCAAGATATAA 1237
seq_name: /SIDS8/gcgdata/geneseq/geneseq/NAL986.DAT: AAN60566
seq_documentation_block:
ID AAN60566 standard; DNA; 1367 BP.
XX
AC AAN60566;
XX
DT 22-AUG-1991 (first entry)
XX
DE Sequence encoding human pregastric lipase.
XX
KW Cystic fibrosis therapy; enzyme; lipase deficiency; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 47..103
FT /tag= a
FT mat_peptide 104..1243
FT /tag= b
XX
PN WO8601532-A.
XX
PD 13-MAR-1986.
XX
PF 15-AUG-1985; 85WO-G800364.
XX
PR 21-AUG-1984; 84GB-0021210.
PR 15-AUG-1985; 85WO-G800364.
PR 01-JAN-1986; 86GB-0008897.
XX
PA (CELL-) CELTECH LTD.
PA (LOWE/) LOWE P A.
XX
PI Lowe PA;
XX
DR WPI; 1986-081634/12.
DR P-PSDB; AAP60658.
XX
PT New gastric lipase protein, esp. of human origin - for treating
PT lipase deficiency, and DNA sequences coding for it
XX
PS Disclosure; Fig 3; 39pp; English.
XX
CC The inventors claim a pregastric lipase protein and a gene encoding
CC it. Gastric lipase protein is useful for oral administration to
CC treat lipase deficiency, e.g. cystic fibrosis or pancreatitis.
XX
SQ Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

alignment_scores:
Quality: 1124.00 Length: 405
Ratio: 3.416 Gaps: 2
Percent Similarity: 81.235 Percent Identity: 52.099
alignment_block:
US-09-333-159-47 x AAN60566 ..
Align seg 1/1 to: AAN60566 from: 1 to: 1367

14 ArgMetGluMetTyrPheLeuLeuValAlaTyrMetPheGlnArgAs 30
|||
38 AGTCCAAATGTGGCTGCTTTTACAAATGGCAAGTTTGATA..... 79
30 nValAsnSerValHisMetProThrLysAlaVal..... 41
|||
80 .....TCTGCTACTGGGGACTACATGGTTTGTGGAAATTCATC 122
|||
42 .....AspProGluAlaPheMetAsnIleSerGluIleIleGlnHisGln 56
|||

```

```

123 CTGGAAGCCCTGAAGTACTATGAACATTAGTCAGATGATTACTATTGG 172
57 GlyTyrProCysGluGluTyrGluValAlaThrGluAspGlyTyrIleLe 73
173 GGATACCAATGAAGATATGAAGTTGTGACTGAAGATGGTTATATTCT 222
73 userValAsnArgIleProArgGlyLeuValGlnProLysLysThrGlyS 90
223 TGAAGTCAATAGATTCTTATGGGAAGAAAAATTCAGGGAATACAGGCC 272
90 erArgProValValLeuLeuGlnHisGlyLeuValGlyGlyAlaSerAsn 106
273 AGAGACCTGTTGTGTTTGGCAGTGTGTTGCTTCATCAGCCCAAAAC 322
107 TrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAl 123
323 TGGATTTCACACCTCCCAACAAACAGCCTTGCCTTCATTCGGCAGATGC 372
123 aGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerArgL 140
373 TGGTTATGATGTGGCTGGCGACAGCAGAGGAGAAACACCTGGGCCAAA 422
140 ysHisLysThrLeuSerLeuAspGlnAspGluPheTrpAlaPheSerTyr 156
423 GAAACTTGTACTATTACACAGATTCCAGTTGAATTCCTGGCTTTCAGCTT 472
157 AspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLeuG 173
473 GATGAATGGCTTAATATGACCTTCCAGCCACAAATCGACTTCATGTAAA 522
173 nLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSerGlnGlyThr 190
523 GAAAACTGGACAGACAGAGCTACACTATGTGGCCATCCCGAGGCCACA 572
190 hrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLysIle 206
573 CCATTGGTTTATTCCTTTCCACCAATCCAGCCTGGCTAAAGAAATC 622
207 LysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLysSe 223
623 AAAACCTTCTATGCTCTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 672
223 rProGlyThrLysPheLeuLeuProAspMetMetIleLysGlyLeuP 240
673 CCTTAAACAAATAGATTGTTCTCAATCCCTCTCAAGTTTATAT 722
240 heGlyLysGluPheLeuTyrGlnThrArgPheLeuArgGlnLeuVal 256
723 TTGGTGACAAATATCTTACCACACACTCTTTGATCAATTTCTGCT 772
257 IleTyrLeuCysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMe 273
773 ACTGAAGTGTGCTCGCTGAGATGCTGAATCTCCTTTGCAGCAATGCCTT 822
273 tLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetSerArgAlaAs 290
823 ATTTATAATTTGGGATTTGACAGTAAAGCTTTAACACAGAGTCGCTGG 872
290 erValTyrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeu 306
873 ATGTGTATCTATCATATAATCCAGCAGGAACTTCTGTTCAAAACATGTC 922
307 HisTrpSerGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpG 323
923 CATTTGACCCAGCGCTGTAAAGTCTGGGAATTCGAAGCTTATGACTGGG 972
323 ySerGluThrLysAsnLeuGluLysCysAsnGlnProThrProValArgT 340
973 AAGCCCACTTCAGAAATAGATGCTACTATGATCAGTCCCAACCTCCCTACT 1022
340 yrArgValArgAspMetThrValProThrAlaMetTrpThrGlyGlyGln 356
1023 ACAATGTACACGCCATGAATGACCAATTCAGTGTGAGTGTGAGCGGTGCAAG 1072

```

```

357 AspTrpLeuSerAspProGluAspValLysMetLeuLeuSerGluValTh 373
1073 GACCTGTGGCTGACCCCAAGATGTTGGCCTTTTGGCTTCCAAACTCCC 1122
373 rAsnLeuIleTyrHisLysAsnIleProGluTrpAlaHisValAspPheI 390
1123 CAATCTTATTTACCACAAGGAGATTCCITTTTACAATCACTTGGACTTTA 1172
390 leTrpGlyLeuAspAlaProHisArgMetTyrAsnGluIleIleHisLeu 406
1173 TCTGGCCAATGATGCCCTCAAGAAGTTTACAATGACATGCTTCTCTATG 1222
407 MetGlnGlnGluGlu 411
1223 ATATCAGAAGATAAA 1237

```

seq_name: /SIDS8/gcgdata/geneseq/geneseq/NA1996.DAT:AAT58916

seq_documentation_block:

ID AAT58916 standard; DNA: 1367 BP.

XX AC AAT58916;

XX DT 19-AUG-1997 (first entry)

XX DE Human gastric lipase coding sequence.

XX KW Duodenal; gastric; lipase; transgenic; plant; recombinant; extract; ss;
 KW food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;
 KW hydrolysis; trans-esterification; substrate; enzyme; biofuel; human.

XX OS Homo sapiens.

Key	Location/Qualifiers
CDS	47..1243
FT	/*tag= a
FT	/product= gastric lipase
FT	47..103
FT	/*tag= b
FT	mat_peptide 104..1240
FT	/*tag= c

XX PN W09633277-A2.

XX PD 24-OCT-1996.

XX PF 19-APR-1996; 96WO-FR00606.

XX PR 20-APR-1995; 95FR-0004754.

XX PA (BIOC-) BIOCEM SA.

XX PA (LJOU) INST RECH JOUVEINAL.

XX PI Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;

XX PI Merot B;

XX DR WPI; 1996-485783/48.

XX DR P-PSDB; AAW09383.

XX CC This is the nucleotide sequence encoding the human pre-duodenal (i.e.
 CC gastric) lipase enzyme. The sequence can be used to generate transgenic
 CC plants producing recombinant lipase in an enzymatically active form.
 CC Alternately bases encoding amino acids 20-23 or 20-73 (1-4 or 1-54 of
 CC the mature protein, respectively) can be deleted to form the derivatives
 CC designated delta-4 or delta-54 respectively. Plants, or their extracts,
 CC expressing the lipases or the truncated derivatives, can be used:

XX PS Claim 3; Fig 4; 130pp; French.

CC (a) as pharmaceuticals or food to facilitate absorption of fat, either
 CC in healthy subjects or in patients with inadequate levels of
 CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 CC pancreatic insufficiency, in the elderly or in patients undergoing medical
 CC treatment which alters fat adsorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.
 XX
 SQ Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

alignment_scores:
 Quality: 1124.00 Length: 405
 Ratio: 3.416 Gaps: 2
 Percent Similarity: 81.235 Percent Identity: 52.099

alignment_block:

US-09-333-159-47 x AAT58916

Align seg 1/1 to: AAT58916 from: 1 to: 1367

14 ArgMetGluMetTrpLeuLeuLeuValAlaValMetPheGlnArgAs 30
 38 AGGTCCAAATGTGGCTTTTAAACAAATGGCAAGTTTGATA..... 79
 30 nValAsnSerValHisMetProThrLysAlaVal..... 41
 80TCGTACTGGGGACTACATGTTTGTGGAAATATACATC 122
 42AspProGluAlaPheMetAsnIleSerGluIleGlnHisGln 56
 123 CTGGAAGCCCTGAAGTACTATGAACATTAAGTACATGATTACTATTGG 172
 57 GlyTyrProCysGluGluTyrGluValAlaThrGluAspGlyTyrIleLe 73
 173 GGATACCAATGAAGATATGAAGTTGTGACTGAAGATGTTATATCT 222
 73 uSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrGlyS 90
 223 TGAAGTCAATAGATTCTTATGGGAGAGAAATTCAGGNAATACAGGCC 272
 90 erArgProValValLeuLeuGlnHisGlyLeuValGlyAlaSerAsn 106
 273 AGAGACCTGTGTGTTTTCAGCATGGTTGCTTGTCATCAGCCACAAAC 322
 107 TrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAl 123
 323 TGGATTTCCAACTCCGCGAACACAGCCTTGCCTTCATCTGGCAGATGC 372
 123 aglyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerArgL 140
 373 TGGTTATGATGTGTGGCTGGGCAACAGCAGAGGAAACACCTGGGCCAGA 422
 140 ysHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTyr 156
 423 GAAACTGTGTACTATTACCAAGATTACAGTTGAATTCCTGGCTTTCAGCTT 472
 157 AspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLeuG 173
 473 GATGAATGGCTTAATATGACCTTCCAGCCACAAATCGACTTCATGTGAAA 522
 173 nLysThrGlyGlnGluLysIleTyrValGlyTyrSerGlnGlyThr 190
 523 GAAACTGGACAGAGCAGCAGTACTATGTGGCCATTCACAGGGCACCA 572
 190 hrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLysIle 206
 573 CCATTTGGTTTATTGCTTTTCCACCAATCCAGCCTGGCTTAAAGAAATC 622
 207 LysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLysSe 223
 ||||||||

623 AAAACCTTCTATGCTCTAGCTCTGTTGCCACTGTGAAGTATACAAAAG 672
 223 rProGlyThrLysPheLeuLeuProAspMetMetIleLysGlyLeup 240
 673 CCTTATTAACAACTTATGTTCTCAATCCCTCTCAAGTTATAT 722
 240 heGlyLysLysGluPheLeuTyrGlnThrArgPheLeuArgGlnLeuVal 256
 723 TTGGTGACAAAATATTCTACCCACACAACTTCTTTGATCAATTTCTTGC 772
 257 IleTyrLeuCysGlyGlnValIleLeuAspGlnIleCysSerAsnIle 273
 773 ACTGAAGTGTGCTCCCTGGAGATCTCAATCTCTTTGCACCAATGCCT 822
 273 tLeuLeuGlyGlyPheAsnThrAsnAsnMetSerArgAlas 290
 823 ATTATATTTTGGATTGTGACAGTAAAGACTTTAACACGAGTCTGCTGG 872
 290 erValTyrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeu 306
 873 ATGTGTATCTATCATATAATCCAGCAGGAACCTCTGTTCAAAACATGTC 922
 307 HisTrpSerGlnAlaValAsnSerGlyLeuLeuArgAlaPheAspTrpG 323
 923 CATTTGGACCCAGGCTGTAAAGTCTGGGAAATTTCAAGCTTATGACTGGG 972
 323 ySerGluThrLysAsnLeuGluLysCysAsnGlnProThrProValArgT 340
 973 AGCCGAGTTTCAGATAGGATGACATGATGATGATGATGATGATGATG 1022
 340 yrArgValArgAspMetThrValProThrAlaMetTrpThrGlyGlyGln 356
 1023 ACAATGTGACAGCCATGAATGTACCAATTTGAGTGTGAACGGTGGCAAG 1072
 357 AspTrpLeuSerAsnProGluAspValLysMetLeuLeuSerGluValTh 373
 1073 GACCTGTGGCTGACCCCAAGATGTGGCCTTTTGTCTCCAAACTCCC 1122
 373 rAsnLeuIleTyrHisLysAsnIleProGluTrpAlaHisValAspPheI 390
 1123 CAATCTTATTACCAAGGAGATTCCTTTTACAACTCAGCTGGACTTTA 1172
 390 leTrpGlyLeuAspAlaProHisArgMetTyrAsnGluIleIleHisLeu 406
 1173 TCTGGCAATGGATGCCCTCAAGAAGTTTACAAATGACATTTGTTCTATG 1222
 407 MetGlnGlnGluGlu 411
 1223 ATATCAGAAGATAAA 1237
 seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF28689
 seq_documentation_block:
 ID: AAF28689 standard; cDNA; 1308 BP.
 XX
 AC: AAF28689;
 XX
 DT: 05-APR-2001 (first entry)
 XX
 DE: Human protein HP03372 coding sequence #2.
 XX
 KW: Human; hydrophobic domain; immune deficiency; autoimmune disorder;
 KW: allergy; tissue growth; regeneration; wound healing; burn; tumour;
 KW: periodontal disease; thrombolytic condition; haemostatic condition;
 KW: infection; ss.
 XX
 OS: Homo sapiens.
 XX
 PN: WO200102563-A2.
 XX
 PD: 11-JAN-2001.
 XX
 PF: 16-JUN-2000; 2000WO-JP03943.

alignment_block:

US-09-333-159-47 x AA234958 ..
Align seg 1/1 to: AA234958 from: 1 to: 1718

1 MetLeuGluThrLeuSerArgGlnTrpIleValSerHisArgMetGluMe 17
232 GTCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 281
17 tTrpLeuIleLeuValAlaTyrMetPheGlnArgAsnValAsnSerV 34
282 A.....GCAGAAACATCAAC..... 297
34 alHisMetProThrLysAlaValAspProGluAlaPhe...MetAsnIle 49
298CCTTCAGTGTATGTCATATGTGCC 321
50 SerGluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGluValAl 66
322 TCTTCTCTCATTTGTCAT...GGATACAAAGTGTCAAGAACACGAGTTAC 368
66 aThrGluAspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuV 83
369 AACTGATGATGGTTACATCTGACCTGCAAGAGATCCACAGAGGTCGAG 418
83 alGlnProLysLysThrGlySerArg.....ProValValLeuLeuGln 97
419 GTAAACAGCAGTGGAGTGGCAGCAAGGAACCAACAGTGGT...ATACAA 465
98 HisGlyLeuValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAs 114
466 CATGGAGTCTTTGTAGATGTATGACATGCTTCTAAACCCACACGAGCA 515
114 nSerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyA 131
516 AGATCTCCCTGTTAGTTTACCTGATATGATGATTTGACGTGTGGATTGCA 565
131 snSerArgGlyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAsp 147
566 ACACAGAGAAACACAGATATAGTCGCGACACATCTCATTTGGACCCCTCT 615
148 GlnAspGluPheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLe 164
616 AGCCAGCCCTATTGGAATTTGGTGTGGGATGAACCTGCTCCCTATGATTT 665
164 uProAlaValIleAsnPheIleLeuGlnLysThrGlyGlnGluLysIleT 181
666 CCCTGGCGTGTAAATATGTTATGTTACAGCAACAGGGGCGAG...AAGATCA 712
181 yTyrValGlyTyrSerGlnGlyThrThrMetGlyPheIleAlaPheSer 197
713 ATTACGTTGGCCATTTCATTGGGAACCTTTGGTAGCTTTGGCATCCTCTCG 762
198 ThrMetProGluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaPr 214
763 GAA...GGAATATGGTTACCCAGTGAATACAGCAGCCCTGTTGAGCCC 809
214 oileAlaThrValLysHisAlaLysSerProGlyThrLysPheLeuLeuL 231
810 TATAGCCTATTATAGCCACATGAATACA.....GCACCTGGTGTCTGTTG 853
231 euProAspMetMet.....IleLysGlyLeuPheGlyLysLysGlu 244
854 CACCCAAAGTCCTTTGTTGGTGAGATCACTACCCTCTTCGCTAGCAGAA 903
245 Phe.....LeuTyrGlnThrArgPheLeuArgGlnLeuValII 257
904 TTTAATCCAAAGGGTTAGCTGTTGATGCTTCTCAAGTCT..... 945
257 eTyrLeuCysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMetL 274
946CTCTGTGCTACCCCTGGGATAGAC.....TGCTATGACTGTGTGA 985

274 euLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlaSer 290
986 CTCGACTAACTGGTAAAAATTGC...TGCCTCAATTCTTCAACTGTTGAT 1032
291 ValTyrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuH1 307
1033 CTATTCCTGATGATGAGCCTCAGCTCAACATCAACAAGAACATGGTGCA 1082
307 sTrpSerGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlyS 324
1083 CTGGGCTCAGACTGTAGACTTGGGGGCTTGACAAAAATTCAAATATGTGA 1132
324 erGluThrLysAsnLeuGluLysCysAsnGlnProThrProValArgTyr 340
1133 GACGAGACTATAACATTATGACTATGAGAAATATTTCTCCAATCTAT 1182
341 ArgValArgAspMet.....ThrValProThrAlaMetTrpThrGlyG1 355
1183 AACCTTTCACACATCCCCACGATCTCCCTCTCTTCATTAGCTATGGTGG 1232
355 yGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLeuSerGluV 372
1233 AAGAGATGCACTTTCAGATGCTCGTGATGTGAGAATTTGCTTGAATAA. 1281
372 alThrAsnLeuIleTyrHisLysAsn.....Ile 381
1282CTCAAGTTCATGATGAGAAACAAGCGCAGCGTTCAGTTCATC 1323
382 ProGluTrpAlaHisValAspPheIleTrpGlyLeuAspAlaProHisAr 398
1324 CAGGATATGCTCATGCTGACTACATTATGGGTTCATGCAATGCCAAGGACTT 1373
398 gMetTyrAsnGluIleIleHisLeuMetGlnGlnGlu 410
1374 GGTGTATATGCTGCTTTTCATTTTCAATCATCAA 1410

seq_name: /SIDSE/gcgdata/geneseq/geneseq/NA2000.DAT:AA234956

seq_documentation_block:

ID AA234956 standard; cdna; 1483 BP.
XX AA234956;
AC AA234956;
XX XX
DT 28-FEB-2000 (first entry)
XX
DE Rice acid triacylglycerol lipase cdna.
XX
KW Triacylglycerol lipase; rice; fatty acid; seed oil;
KW vegetable oil; transgenic plant; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 32..1264
FT /*tag= a
XX
PN WO9955883-A2.
XX
PD 04-NOV-1995.
XX
PF 29-APR-1995; 99WO-US09280.
XX
PR 30-APR-1995; 98US-0083688.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon EB, Cahoon RE, Kinney AJ, Rafalski JA;
XX
DR WPI; 2000-062036/05.
XX
XX P-PSDB; AAY32307.
PT Novel plant triacylglycerol lipase polynucleotides used to alter the

PT level of the enzyme in transgenic plants -

XX Claim 2: Page 44: 65pp: English.

CC This is the nucleotide sequence of the cDNA insert in clone
CC rlr72.pk0015.b2 encoding the entire rice acid triacylglycerol
CC lipase (TAGL) (see AAY32307). The clone was isolated from a rice
CC leaf cDNA library. Novel acid and neutral TAGL polypeptides
CC (see AAY32301-17) and polynucleotides (see AAZ34950-66) from corn,
CC Catalpa, rice, soybean and wheat tissues are disclosed. The enzymes
CC may be prepared recombinantly and used to raise antibodies, which
CC are used for detecting the enzymes in situ in cells or in vitro in
CC cell extracts. The polynucleotides may be used to create transgenic
CC plants in which the TAGL levels are present at higher or lower levels
CC than normal, or in cell types or developmental processes where they are
CC not normally found. This would alter the level of triacylglycerol and
CC cholesterol esters found in those cells. Accumulation of fatty acids
CC with unusual structures may be a positive phenotype in plants used for
CC foods. In addition, it may be desirable to eliminate expression of TAGL
CC genes for certain applications. TAGL enzymes may also be useful for the
CC processing of plant seed oils and for the development of novel seed
CC oils. The TAGL enzymes can also be used as targets to facilitate the
CC design and/or identification of inhibitors of those enzymes that may be
CC useful as herbicides. This is desirable because inhibition of the
CC activity of either of the enzymes could lead to an inhibition of plant
CC growth. The polynucleotides also serve as a source of probes and
CC primers, which are useful for genetic mapping, as markers for traits
CC linked to those genes, and to isolate homologous sequences from other
CC species.

XX Sequence 1483 BP; 375 A; 334 C; 344 G; 430 T; 0 other;

alignment_scores:

Quality: 502.00 Length: 365
Ratio: 2.109 Gaps: 13
Percent Similarity: 65.205 Percent Identity: 33.973

alignment_block:

US-09-333-159-47 x AAZ34956

Align seg 1/1 to: AAZ34956 from: 1 to: 1483

57 GlyTyrProCysGluGluTyrGluValAlaThrGluAspGlyTyrIleLe 73
194 GGCTACCCCTCCACGAGCACACAGCTTGAACAAACAAAGATGGATCCCTTT 243
73 uSerValAsnArgIleProArgGlyLeuValGln...ProLysLysThrG 89
244 ATCTCTTCAGATATCCACATGCGCAAAATAAGCAGCAGATAGTACTG 293
89 lySerArgProValLeuLeuGlnHisGlyLeuValGlyAlaSer 105
294 GC.....CCTCCAGCTTTTCTTCAACATGCTCTTTTCAGGGAGGAGAC 337
106 AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
338 ACATGGTTTCATAAACTCGTGACCAATCACTTGGGTATATCTCTGTGTA 387
122 pAlaGlyPheAspValTrpMetGlyValAsnSerArgGlyValAlaTrpSerA 139
388 TAACGGTTTTCATGTTGGATGGGATGTCCTGGGAACGCGTTGGAGTA 437
139 rgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
438 AAGGTCATCAACCTTTTCTCTCATGATAAGCTTTTCTGGGATGGAGC 487
156 TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
488 TGGCAAGAGTAGTGAATAGTACCTTTTAGCAATGTAGCTATGTGTA 537
172 uGlnLysThrGlyGlnGluLysIleTyrValGlyTyrSerGlnGlyT 189

538 TACAGTCACA...CAGTCCAAATTCATATATGTGGGGCATTACACAGGAA 584
189 hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
585 CTATAATGGGTTTGGCGCTTTG...ACGATGCCGAATAGTAAATAG 631
206 IleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLys 222
632 ATTACGCTCGCAGCAGCTCTTTGCTCTATTCTTATCTTGTATGATGTTAG 681
222 sSerProGlyThrLysPheLeuLeu.....LeuProA 233
682 TGCT.....AGTTTGTCTCAGACAGCAGTCGCCCATGCTCTTGATC 722
233 spMetMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThr 249
723 AGATGCTTGT.....ACTATGGGAATTCACAGCTGAACCTCGGTAGC 766
250 ArgPheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAs 266
767 GACATGGGGTTCAAATAGTAGATCTTTTGGCATGGTGAACAGCTGGA 816
266 pGlnIleCysSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnA 283
817 T.....TGCAACAATTTGCTATCTCGCATACAGGGGAAACTGT...T 857
283 snMetAsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGly 299
858 GCTTCAATACATCAAGGATGATTATTATTGGAGTATGAACCTCATCCA 907
300 ThrSerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyG 316
908 TCATCGCAAAAAATCTGCACCATCTTTTTCAGATGATCAGGAAGGAC 957
316 uLeuArgAlaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysA 333
958 TTTCCGCAAGTATGACTATGGTTA...TTGGGAAACCTAAGGCGCTAGC 1004
333 snGlnProThrProValArgTyrArgValArgAspMet.....ThrVal 347
1005 GTCATTTGCGTCTCCCGCATTTGACCTTAAGCAGCATACCAAGATCACTG 1054
348 ProThrAlaMetTrpThrGlyGlyGlnAspTrpLeuSerAsnProGluAs 364
1055 CCCATATGGGATGGGATGGAGGCTTTCATTCATGGCTGATGTAACCGA 1104
364 pValLysMetLeuLeuSerGlyVal.....ThrAsnLeuIleTyrH 378
1105 TGTTCAAGCTATATCAGAGAGCTGGGATCTACACCAAGAACTTCTGTAC 1153
378 lsLysAsnIleProGluTrpAlaHisValAspPheIleTrpGlyLeuAsp 394
1154ATTGGTCACTATGCCCATATTGATTTTGTATGAGCGTGAAG 1195
395 AlaProHisArgMetTyrAsnGluIleHisLeuMetGlnGln 409
1196 CGGAAAGATGATGTTTATGTGGACCTTAATAAGATTTCTTAGGGAA 1240

OM of: US-09-333-159-47 to: Issued_Patents_NA:* out_format : pfs
 Date: Sep 1, 2001 8:18 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEI=frame2.p2n.model -DEV=xlp
 -Q/cgn2_1/USPTO_spool/US09333159/runat_31082001_122330_1399/app_query.fasta_1.488
 -DB=issued_Patents_NA -OFMT=fastap -SUFFIX=p2n.rni
 -GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000
 -LOOPEXT=0.000 -GAPEXT=4.500 -GAPEXT=0.050 -GAPEXT=10.000
 -GAPEXT=0.500 -GAPEXT=6.000 -GAPEXT=7.000 -GAPEXT=10.000
 -GAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
 -MATRIX=blossum62 -TRANS=human40.cdi -LIST=45 -DOALIGN=200
 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=200000000
 -USER=US09333159_cgn1_1_62 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:

Query: US-09-333-159-47
 Query length: 423
 Database: Issued_Patents_NA:*
 Database sequences: 324599
 Database length: 9465562
 Search time (sec): 69.100000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-227-108-2	1141.50	2461.52	1.7e-129	1137	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-09-073-674-2	1141.50	2461.52	1.7e-129	1137	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-227-108-4	1141.50	2461.49	1.7e-129	1140	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-073-674-4	1141.50	2461.49	1.7e-129	1140	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-6	1141.50	2461.42	1.7e-129	1146	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-073-674-6	1141.50	2461.42	1.7e-129	1146	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-1	1141.50	2457.71	2.8e-129	1528	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-073-674-1	1141.50	2457.71	2.8e-129	1528	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-356-952-12	271.50	510.09	8.5e-21	43676	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-751-782-2	155.00	326.10	1.5e-10	178	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-925-171-2	155.00	326.10	1.5e-10	178	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-484-108-21	105.00	187.71	0.0077	1676	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-484-108-21	105.00	187.71	0.0077	1676	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-602-359A-24	102.00	187.29	0.0081	1041	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-232-519-1	98.50	178.16	0.0262	1167	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-456-956-1	98.50	178.16	0.0262	1167	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-115-446-1	92.00	153.99	0.5820	2521	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-115-446-5	92.00	153.96	0.5839	2526	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-471-119A-1	92.00	116.26	73.51	46899	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-080-386-3	91.00	153.79	0.5967	2160	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-742-185-102	91.00	153.79	0.5967	2160	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-968-543C-1	89.00	136.17	5.72	6027	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-755-559-2	85.00	140.54	3.27	2180	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-210-474-2	85.00	140.54	3.27	2180	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-602-359A-30	84.50	149.29	1.06	1017	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-759-945-1	84.50	137.05	5.11	2625	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-07-707-367-1	84.50	133.78	7.78	3383	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-386-607-1	84.00	138.47	4.26	2160	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-07-882-329-4	82.50	141.46	2.90	1329	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-183-213-4	82.50	141.46	2.90	1329	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-446-803-4	82.50	140.29	3.37	1455	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-861-837-4	82.50	140.29	3.37	1455	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-600-656-4	82.50	140.29	3.37	1455	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-170-670-9	82.50	140.29	3.37	1455	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-170-670-14	82.50	140.29	3.37	1455	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-193-068-9	82.50	140.29	3.37	1455	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-193-068-13	82.50	140.29	3.37	1455	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-183-412-9	82.50	140.29	3.37	1455	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-183-412-13	82.50	140.29	3.37	1455	

/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-600-908A-11 + 82.50 140.26 3.39
 /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-683-838A-11 + 82.50 140.26 3.39
 /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-483-101-1 + 82.50 122.45 33.26
 /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-348-891A-1 + 82.50 109.43 176.52
 /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-905-817-1 + 82.50 109.43 176.52
 /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-2

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-2

seq_documentation_block:
 ; Sequence 2, Application US/08227108
 ; Patent No. 5807726

; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/227,108
 ; FILING DATE: 03-APR-1994

; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fanucci, Allan A.
 ; REGISTRATION NUMBER: 30,256
 ; REFERENCE/DOCKET NUMBER: 7620-033
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 869-8864/9741
 ; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1137 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1137
 ; US-08-227-108-2

alignment_scores:
 Quality: 1141.50 Length: 372
 Ratio: 3.624 Gaps: 3
 Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
 US-09-333-159-47 x US-08-227-108-2

Align seg 1/1 to: US-08-227-108-2 from: 1 to: 1137

37 ProThrLysAlaValaspProGluAlaPheMetAsnIleSerGluIle1 53
 |||||
 19 CCCACA.....AACCCCTGAGTGACCATGAATTAAGTCAGATGAT 59
 53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspG 70
 |||||
 60 CACCTACTGGGATACCCAGCTGAGGAATATGAAGTTGTGACCGAAGACG 109
 70 lYTyrlleuSerValAsnArgIleProArgGlyLeuValGlnProLys 86

|||||.....:||||||| |||
110 GTTATATCCCTGGGATCGACAGAAATTCCTTATGGGAGAAAAAATTTCAGAG 159
:
87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGlyGL 103
::: |||:|||||
160 AATATAGCCGGAGACCTGTGTGCATTTTGCAACACGGTTGCTCGCATC 209
:
103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
:||||:.....:|||||
210 AGCCACAATACTGGATCTCCAACCTGCCCAACACAGACTGGCCCTTCATCC 259
:
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
|||||.....:|||||
260 TGGCCGACGGCGGTACGAGCTGTGGCTGGGGACACAGCGGGGCACACC 309
:
137 TrpSerArgLyshisLeysThrLeuSerIleAspGlnAspGluPheTrpAl 153
|||||.....:|||||
310 TGGGCCAGGAGGAATCTGTACTACTCCCGCGACTCCGTGCANATTCCTGGC 359
:
153 apHeSerTyraSpGluMetAlaArgPheAspLeuProAlaValIleAsnP 170
|||||.....:|||||
360 TTTcAGCTTTGACGAGATGGCTAAATATGACCTTCCCGCCACCATTCGACT 409
:
170 heIleLeuGlnLysThrGlyGlnGluylysIleTyTYrValGlyTySer 186
|||||.....:|||||
410 TCATCTTTGAAGAAAACGGGACAGACAGAAGTACACTACGTTGGCCATTC 459
:
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
|||||.....:|||||
460 CAGGGCACCACTATGGTTTCATCGCCTTTTCCACCAATCCCACAGCTGGC 509
:
203 aGlnLysIleLysMetTyrrPheAlaLeuAlaProIleAlaThrVallysh 220
|||||.....:|||||
510 GAACAGGATCAAAACCTCTCATGCAATTAGCTCCCGTTGCCACCGTGAAGT 559
:
220 IsAlaLysSerProGlyThyLysPheLeuLeuLeuProAspMetMetIle 236
|||||.....:|||||
560 ACACCGAACCCTGTTAAACAACACTCATGTCTGTCCTCTCGTCTCCTCTC 609
:
237 LysGlyLeuPheGlyLysLysGluPheLeuTyrglnThrArgPheLeuAr 253
|||||.....:|||||
610 AAGCTTATATTGGAAACAAAATATTC...TACCCACACCACTCTTTGA 656
:
253 gGln...LeuValIleTyrrLeuCysGlyGlnValIleLeuAspGlnIleC 269
|||||.....:|||||
657 TCATATTTCTGCCACCGAGGTATGCTCCCGGACAGCGTGGATCTCCTCT 706
:
269 ysSerAsnIleMetLeuLeuGlyClyPheAsnThrAsnAsnMetAsn 285
|||||.....:|||||
707 GCAGCAAACGCCCTGTTTATCATTTGTGGATTGACACTATGAACCTTGAC 756
:
286 MetSerArgAlaSerValTyrrAlaAlaHisThrLeuAlaGlyThrserva 302
|||||.....:|||||
757 ATGAGTCGCTTGGATGTGTATCTGTACAAATACCACAGAACATCCGCT 806
:
302 lGlnAsnileLeuHistrpSerGlnAlaValAsnSerSerglyLeuArgA 319
|||||.....:|||||
807 TCAGAACGTGCTCCACCTGGTCCCGAGGTGTTAAGTCTGGGAAGTCTCCAG 856
:
319 laPheAspTrpGlySerGluThyLysAsnLeuGluLysCysAsnGlnPro 335
|||||.....:|||||
857 CTTTTGACTGGGAAGCCCCAGTTTCAGAACATGATGCATCTATCATCAGAGC 906
:
336 ThrProValArgTyArgValArgAspMetThrValProthrAlaMetTr 352
||| |.....:|||||
907 ATGCCTCCCTACTACAACCTGCACAGACATGCATGTGCCATCGCAGTGTG 956
:
352 pThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValIyleMetLeuL 369
|||||.....:|||||
957 GAACGGTGGCAACGACTTGTCTGCGCCGACCTTCACGATGTTGACCTTTTGC 1006
:
369 euSerGluValThrAsnLeuIleTyrrHisLysAsnIleProGluTrpAla 385
|||||.....:|||||

```

1007 TTTCCAGCTCCCAATCATTTACACAGGAAGATTCCTCCTTACAAT 1058
386 HisValaspPheiletrpGlyLeuAspAlaProHisArgMetTyrAsnGI 402
|||||.....|.....|.....|.....|.....|.....|.....|.....|
1057 CACTTGGACTTATCTGGCCATGGATGCCCTCAAGCGGTTTACAATGA 1106
402 ullelleHisLeuMet 407
|||||.....|.....|
1107 AATTGTTCATGATG 1122

seq_name: /cgn2_6/ptodata/1/lina/5B_COMB.seq:US-09-073-674-2

seq_documentation_block:
; Sequence 2, Application US/09073674
; Patent No. 5958189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junten, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1137 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1137
; US-09-073-674-2

alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
US-09-333-159-47 x US-09-073-674-2 ..

Align seg 1/1 to: US-09-073-674-2 from: 1 to: 1137

37 ProThrAlaValaspProGluAlaPheMetAsnIleSerGluIleI 53
|||||.....|.....|.....|.....|.....|.....|.....|.....|
19 CCCACA.....AACCTCGAAGTGACCATGAATATAGTCAGATGAT 59

53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAsp 70

```



```

910  ATGCGCTCCTACTACAACTGCAGACATGTCATGTGCCAATGCCAGTGTG 959
352  pThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
960  GAACGGTGGCAACGACTTCTGGCGCACCTTCACGATGTTGACCTTTGC 1009
369  euSerGluValThrAsnLeuLeuTyrHisLysAsnIleProGluTrpAla 385
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1010  TTTCAAGCTCCCAATCTCATTTACACAGGAAGATTCTCTCTTCAAT 1059
386  HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnG 402
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1060  CACTTGGACTTATCTGGCCATGGATGCCCTCAAGCGGTTTACAATGA 1109
402  uIleIleHisLeuMet 407
      |||:|||||:|||||
1110  AATTGTTTCCATGATG 1125

seq_name: /cgn2_6/ptodata/1/1na/5B_COMB.seq:US-09-073-674-4

seq_documentation_block:
; Sequence 4, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1140
US-09-073-674-4

alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
US-09-333-159-47 x US-09-073-674-4

```

Align seg 1/1 to: US-09-073-674-4 from: 1 to: 1140

```
37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIle1 53
   |||||
   :|||:
   :|||:
22 CCCACA.....AACCTGAAGTGACCATGAATATAAGTCAGATGAT 62
   |||||
   :|||:
   :|||:
53 eelnHISGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspG 70
   |||||
   :|||:
   :|||:
63 CACCTACTGGGGATACCCAGGTGAGGAATATGAAGTTGTGACCGAAGACG 112
   |||||
   :|||:
   :|||:
70 lYtYrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
   |||||
   :|||:
   :|||:
113 GTTATATCTTGGGATCGACAGAAATTCCTTATGGGAGGAGAAATTCAGAG 162
   |||||
   :|||:
   :|||:
87 LysThrGlySerArgProValValLeuLeuGlnHISGlyLeuValGlyG 103
   |||||
   :|||:
   :|||:
163 AATATAGCCGCGAGACCTGTTGCAATTTTGCACACGCGTTGCTCGCATC 212
   |||||
   :|||:
   :|||:
103 ValaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
   |||||
   :|||:
   :|||:
213 ACCCAACAACCTGGATCTCCACCTCCCAACACAGCCTGGCCTTCATCC 262
   |||||
   :|||:
   :|||:
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
   |||||
   :|||:
   :|||:
263 TGGCCGACGCGGGTACGACGTGTGGCTGGGAAACAGCAGGGGCAACACC 312
   |||||
   :|||:
   :|||:
137 TrpSerArgLysHISLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
   |||||
   :|||:
   :|||:
313 TGGGCGAGGAGAACTCTACTACTCTGCGCGGCGGACTCCGTCGAATTCGGGC 362
   |||||
   :|||:
   :|||:
153 aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnP 170
   |||||
   :|||:
   :|||:
363 TTTCAGCTTTCAGGAGATGGCTAAATATGACCTTCCCGCCACCATGACT 412
   |||||
   :|||:
   :|||:
170 heIleLeuGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSer 186
   |||||
   :|||:
   :|||:
413 TCATCTTGAAGAAACCGGACGAGGACAGCAAGCTACACTAGTGGCCATCC 462
   |||||
   :|||:
   :|||:
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
   |||||
   :|||:
   :|||:
463 CAGGCGACCAACCATTTGTTTTCATCGCTTTCCACCAATCCCAAGCTGC 512
   |||||
   :|||:
   :|||:
203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
   |||||
   :|||:
   :|||:
513 GAACGGATCAAAACCTTCTATGCAATAGCTCCCGTGGCCACCGTGAAGT 562
   |||||
   :|||:
   :|||:
220 IsAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetIle 236
   |||||
   :|||:
   :|||:
563 ACACCGAAACCCCTGTTAAACAACATCATGCTCGTCCCTCGTTCCTCITC 612
   |||||
   :|||:
   :|||:
237 LysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArgPheLeuAr 253
   |||||
   :|||:
   :|||:
613 AAGCTTATATTGGAAACAAATATTC...TACCACACCACTCTTTGA 659
   |||||
   :|||:
   :|||:
253 gLIn...LeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
   |||||
   :|||:
   :|||:
660 TCAATTTCTCGCCACCGAGGTATGCTCCCGGAGACGGTGGATCTCCTCT 709
   |||||
   :|||:
   :|||:
269 ySerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsn 285
   |||||
   :|||:
   :|||:
710 GCAGCAACCCCTGTTATCATTTGCGATTGACACATATGAACITGAAC 759
   |||||
   :|||:
   :|||:
286 MetSerArgAlaSerValTyrAlaAlaHISThrLeuAlaGlyThrSerVa 302
   |||||
   :|||:
   :|||:
760 ATGAGTCGCTTGGTGTATCTGTCTACATAATCCAGCAGGAGCAACATCGGT 809
   |||||
   :|||:
   :|||:
302 lGlnAsnIleLeuHISTrpSerGlnAlaValaValaAsnSerGlyGluLeuArgA 319
   |||||
   :|||:
   :|||:
810 TCAGAACGTCCTCCACTGCTCCAGGCTGTTAAGTCTGGGAAGTTCCAAG 859
   |||||
   :|||:
   :|||:
319 laPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
   |||||
   :|||:
   :|||:
```

```
860 CTTTGTACTGGGAGCCAGTTCAGAACATGATGCACATCATCATCAGAGC 909
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
   |||||
   :|||:
   :|||:
910 ATGCTCTCCTACTACAACTGACAGACATGATGTGCCAATCCGAGTGTG 959
   |||||
   :|||:
   :|||:
352 pThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
   |||||
   :|||:
   :|||:
960 GAACGGTGGCAACGACTTGTGGCCGACCTCCAGATGTTGACCTTTTGC 1009
   |||||
   :|||:
   :|||:
369 euSerGluValThrAsnLeuIleTyrHISLysAsnIleProGluTrpAla 385
   |||||
   :|||:
   :|||:
1010 TTTCCAAGCTCCCAATCTCATTTACCAGCAAGATTCCTCCTTACAAT 1059
   |||||
   :|||:
   :|||:
386 HISValAspPheIleTrpGlyLeuAspAlaProHISArgMetTyrAsnG 402
   |||||
   :|||:
   :|||:
1060 CACTTGGACTTATCTGGCCATGATGCCCTCAAGCGGTTTACAATGA 1109
   |||||
   :|||:
   :|||:
402 uIleIleHISLeuMet 407
   |||||
   :|||:
   :|||:
1110 AATGTTTCCATGATG 1125

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-6

seq_documentation_block:
; Sequence 6, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1146 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; US-08-227-108-6
```

```
alignment_scores:
  Quality: 1141.50      Length: 372
  Ratio: 3.624         Gaps: 3
  Percent Similarity: 84.677      Percent Identity: 56.452

alignment_block:
```

US-09-333-159-47 x US-08-227-108-6

Align seg 1/1 to: US-08-227-108-6 from: 1 to: 1146

```
37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIle 53
||||| :||||| :
19 CCCACA.....AACCTTGAGTGACCATGATATAGTCAGATGAT 59
||||| :||||| :
53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspG 70
||||| :||||| :
60 CACCTACTGGGATACCCAGCTGAGGAATATGAAGTTGTGACCGAAGACG 109
||||| :||||| :
70 LysThrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
||||| :||||| :
110 GTTATATCTTGGGATCGACAGAAATCTTATGGGAGGAAATTCAGAG 159
||||| :||||| :
87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGlyG 103
||||| :||||| :
160 AATATAGCCGAGACCTGTTGCAATTTTGGCAACAGGTTGCTCGGCATC 209
||||| :||||| :
103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
||||| :||||| :
210 AGCCACAACTGGATCTCCAACCTGCCCAACACAGCCTGCGCTTCATCC 259
||||| :||||| :
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
||||| :||||| :
260 TGGCGGACGCGGTTACACGCTGCTGCTGGGGAACAGCAGGGGCAACAC 309
||||| :||||| :
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
||||| :||||| :
310 TGGGCGCAGAGGATCTGTACTACTCGCCGACCTCGCTCGAATTCCTGG 359
||||| :||||| :
153 aPheSerTyrAspGluMetalaaArgPheAspLeuProAlaValIleAsn 170
||||| :||||| :
360 TTTCAGCTTTCAGCAGATGGCTAAATATGACCTTCCCGCCACCATGACT 409
||||| :||||| :
170 heIleLeuGlnLysThrGlyGlnGluLysIleTyrValGlyTyrSer 186
||||| :||||| :
410 TCATCTTGAAGAAACGGGACGAGCAGACAGTACACTAGCTGGCCATTC 459
||||| :||||| :
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
||||| :||||| :
460 CAGGCGCACCATGTTGTTTCATCGCTTTCACCAATCCCAAGCTGGC 509
||||| :||||| :
203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
||||| :||||| :
510 GAAACGGATCAAAACCTTCTATGATAGCTCCGCTGCCACCGTGAAGT 559
||||| :||||| :
220 IsAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetile 236
||||| :||||| :
560 ACACCGAAACCCCTGTTAAACAACTCATGCTCGTCCCTTCCTCTTC 609
||||| :||||| :
237 LysGlyLeuPheGlyLysGluPheLeuTyrGlnThrArgPheLeuAr 253
||||| :||||| :
610 AAGCTTATATTGGAACAAATATTC...TACCCACACCACTTCTTGA 656
||||| :||||| :
253 gGln...LeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
||||| :||||| :
657 TCAATTTCTCCCGCCGAGGATGCTCCCGGAGACGGTGTCTCTCT 706
||||| :||||| :
269 ySerAsnIleMetLeuLeuLeuGlyPheAsnThrAsnAsnMetAsn 285
||||| :||||| :
707 GCAGCAACCCCTGTTTATCATTGTTGATTTGACATATGAACCTGAAC 756
||||| :||||| :
286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVa 302
||||| :||||| :
757 ATGAGTCGCTGGATGTGTATCTGTACATAATCCAGCAGGACATCGGT 806
||||| :||||| :
302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArg 319
||||| :||||| :
807 TCAGAACGTGCTCCACTGCTCCAGGCTGTTAAGTCTGGGAAGTTCCAAG 856
||||| :||||| :
319 laPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
||||| :||||| :
```

```
||||| :||||| :
857 CTTTGTGAGTGGGAGCCAGTTCAGAAATGACACTATCATCAGAGC 906
||||| :||||| :
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
||||| :||||| :
907 ATGCTTCCCTACTACACCTGACAGATGATGGCCAATCGCAGTGTG 956
||||| :||||| :
352 pThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
||||| :||||| :
957 GAACGGTGGCAACGACCTGTGTCGGCGGACCCCTCACGATGTTGACCTTTGC 1006
||||| :||||| :
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
||||| :||||| :
1007 TTTCCAAAGCTCCCAATCTCATTTACCACAGGAAGATTCCTCCTTACAAT 1056
||||| :||||| :
386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnG 402
||||| :||||| :
1057 CACTTGACTTTATCTGGCCATGATGCCCTCAAGCGGTTTACAATGA 1106
||||| :||||| :
402 uIleIleHisLeuMet 407
||||| :||||| :
1107 AATTGTTTCCATGATG 1122

seq_name: /csn2.5/ptodata/1/lna/5B_COMB.seq:US-09-073-674-6

seq_documentation_block:
; Sequence 6, Application US/09073674
; Patent No. 5938189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1146 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; US-09-073-674-6
```

alignment_scores:

Quality:	1141.50	Length:	372
Ratio:	3.624	Gaps:	3
Percent Similarity:	84.677	Percent Identity:	56.452

alignment_block:

US-09-333-159-47 x US-09-073-674-6

Align seg 1/1 to: US-09-073-674-6 from: 1 to: 1146

```

37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleL 53
|||||
19 CCACAC.....AACCTGAGAGTGACCATGATATAGTCAGATGAT 59
53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspG 70
|||
60 CACCTACTGGGATACCCAGCTGAGGAATATGAAGTGTGACCGAAGACG 109
70 LysThrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
|||||
110 GTATATCCTTGGGATGACAGAAATTCCTTATGGGAGGAAATTCAGAG 159
87 LysThrGlySerArgProValLeuLeuGlnHisGlyLeuValGlyG 103
|||
160 AATATAGCGGAGACCTGTGTCATTTTGAACACGGTTTGTCTGCGCATC 209
103 ValSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
|||||
210 AGCCACAACTGGATCTCCAACTCCCAACCAACAGCTGGCTTCATCC 259
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
|||||
260 TGGCGGACGCGGTTACGACGTGTGCTGGGAAACAGCAGGCGGCAACACC 309
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
|||||
310 TGGGCGAGAGAAATCTGTACTACTCGCCGACTCCGCGAAATCTGGGC 359
153 aPheSerTrpAspGluMetAlaArgPheAspLeuProAlaValIleAsn 170
|||||
360 TTTCAGCTTTGACGAGATGGCTAAATATGACCTTCCCGCCACCATGACT 409
170 heIleLeuGlnLysThrGlyGlnGluLysIleTyrTrpValGlyTyrSer 186
|||||
410 TCATCTTGAGAAACAGGAGACAGGAACTACACTAGCTGTGGCATTC 459
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
|||||
460 CAGGCGACCACTTGGTTTCTATCCCTTTTCCACCAATCCCAAGCTGCG 509
203 aGlnLysIleLysMetTrpPheAlaLeuAlaProIleAlaThrValLysH 220
|||||
510 GAAACGGATCAAAACCTTCTATGCATAGCTCCCGTGCACCGTGAAGT 559
220 IsAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetile 236
|||||
560 ACACCGAAACCTGTAAACAACTCATGCTCGTCCCTCGTCTCTCTTC 609
237 LysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArgPheLeuAr 253
|||||
610 AAGCTTATATTGGAACAAATAATTC...TACCACACCACTTCTTTGA 656
253 gGln...LeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
|||||
657 TCANATTCGCGCCAGGATAGTTCGCGGAGAGCGGTGGATCTCTCT 706
269 ySerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsn 285
|||||
707 GCAGCAACGCCCTGTTTATCATTTGTGGATTGACATATGAACATTGAAC 756
286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVa 302
|||||
757 ATGAGTCGCTTGGATGTGTATGTGTACATAATCCAGCAGGAACATCGGT 806
302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgA 319
|||||
807 TCAGAACGTGCTCCCACTGTGCTCCAGGCTGTTAAGTCTGGGAAGTCCCAAG 856

```

```

319 laPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
|||||
857 CTTTGTGACTGGGAAGCCAGTTCAGAACATGATGCACACTATCATCAGAGC 906
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
|||||
907 ATGCTCTCCTACTACAACTGACAGACATGCATGTGCCAATCGCATGTG 956
352 pThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
|||||
957 GAACGGTGGCAAGCACTTGTGCGCGACCTCAGCATGTGTGACCTTTGCG 1006
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
|||||
1007 TTTCCAACTCCCAATCTCATTTACCACAGAGATTCCTCTCTTACAAT 1056
386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTrpAsnG 402
|||||
1057 CACTGTGACTTTACTGTGGCCATGGATGCCCTCAAGCGGTTTACAATGA 1106
402 uIleIleHisLeuMet 407
|||||
1107 AATTGTTTCCATGATG 1122

```

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-1

seq_documentation_block:

; Sequence 1, Application US/08227108

; Patent No. 5807726

; GENERAL INFORMATION:

; APPLICANT: Blanchard, Claire

; APPLICANT: Benicourt, Claude

; APPLICANT: Junien, Jean-Louis

; TITLE OF INVENTION: Recombinant Dog Gastric Lipase

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/227,108

; FILING DATE: 03-APR-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Fanucci, Allan A.

; REGISTRATION NUMBER: 30,256

; REFERENCE/DOCKET NUMBER: 7620-033

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 790-9090

; TELEFAX: 212 869-8864/9741

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1528 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

; US-08-227-108-1

alignment_scores:

Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:

US-09-333-159-47 x US-08-227-108-1

Align seg 1/1 to: US-08-227-108-1 from: 1 to: 1528

```
37 ProThrLysAlaValAspProGluAlaPheMetAsnLysSerGluLeuI 53
19 CCCACA.....AACCCCTGAAGTGACCATGAATATAAGTCAGATGAT 59
53 eGlnHisGlnGlyProCysGluGluTyrGluValAlaThrGluAspG 70
60 CACCTACTGGGGAATACCCAGCTGAGGATATGAAGTTGTGACCGAAGC 109
70 LyrTrpLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
110 GTTATATCCTTGGGATCGACAGAATTCCTTATGGGAGGAAAAATTCAG 159
87 LysThrGlySerArgProValValLeuGlnHisGlyLeuValGlyG 103
160 AATATAGCCGCGGAGACCTGTTGCAATTTTGCACACGCTTTCCTCGC 209
103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPhe 120
210 AGCCACAACCTGGATCTCCACCTGCCCAACACAGCGCTGGCCCTTCAT 259
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsn 136
260 TGGCGGACCGCGGTACGACGTGTGGCTGGGGAACAGCAGCGGCAAC 309
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPhe 153
310 TGGGCGGAGGAGATCTGTACTCTCCGCGGACCTCCGTCGAATTCG 359
153 aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIle 170
360 TTTCAGCTTTCACGAGATGGGTAATATGACCTCCGCGCACCATTC 409
170 heLeuGlnLysThrGlyGlnGluLysIleTyrValGlyTyrSer 186
410 TCATCTTGAAGAAACCGGACGACGACGCTACACTACGTTGGCCAT 459
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGlu 203
460 CAGGCGACCATCTGGTTTCATCGCTTTTCCACCAATCCCAAGCTGG 509
203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrVal 220
510 GAAACGGATCAAAACCTTCTATGATAGCTCCGTTGCCACCGTGA 559
220 IsAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetIle 236
560 ACACCGAAACCCCTGTAACAAACATCATGCTCGTCCCTTCCTCT 609
237 LysGlyLeuPheGlyLysGluPheLeuTyrGlnThrArgPheLeu 253
610 AAGCTTATATTGGAAACAAATATTC...TACCCACACCACTTCTTGA 656
253 gGln...LeuValIleTyrLeuCysGlyGlnValIleLeuAspGln 269
657 TCAATTTCTGCCCGGAGGTATGCTCCCGGAGACGGTGGATCTCT 706
269 ysSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnMetAsn 285
707 GCACCAACGCCCTGTTTATCATTTTGTGGATTGACACTATGAAC 756
286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThr 302
757 ATGAGTCCCTGGATGTGTATCTGCATAAATCCAGCAGGAACATCG 806
302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeu 319
807 TCAGAACGTGCTCCACTGGTCCCGAGGCTGTTAAGTCTGGGAAGTCC 856
```

```
319 laPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
857 CTTTGTGACTGGGAAGCCAGTTCAGAACATGATGCACATATCATCAGAGC 906
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMet 352
907 ATGCTCTCCCTACTACAACCTGACAGACATGATGTGCCAATCGCAGTGT 956
352 pThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
957 GAACGGTGGCAACGACTGTGCGCGGACCCCTCAGCATGTTGACCTTTGC 1006
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrp 385
1007 TTTCCAAGCTCCCAATCTCATTTTACCACAGAGATTCCTCTTACAAT 1056
386 HisValaspPheIleTrpGlyLeuaspAlaProHisArgMetTyrAsn 402
1057 CACTTGGACTTATCTGGGCATGATGTCCTCAAGCGGTTTACAATGA 1106
402 uIleIleHisLeuMet 407
1107 AATTGTTCCATGATG 1122
seq_name: /cgn2_6/ptodata/1/lna/5b_COMB.seq:US-09-073-674-1
seq_documentation_block:
; Sequence 1, Application US/09073674
; Patent No. 598189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/073,674
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-DI-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1528 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; US-09-073-674-1
```

alignment_scores:

Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3

Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:

US-09-333-159-47 x US-09-073-674-1

Align seg 1/1 to: US-09-073-674-1 from: 1 to: 1528

```
37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIle1 53
   |||||
19 CCACA.....AACCTGAGTGACCATGAATATAGTCAGATGAT 59
   |||||
53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspG 70
   |||||
60 CACCTACTGGGATACCCAGCTGAGGAATATGAAGTGTGACCGAAGAGC 109
   |||||
70 LysThrLysSerValAsnArgIleProArgGlyLeuValGlnProLys 86
   |||||
110 GTTATATCTCTGGGATCGACAGAAATTCCTTATGGGAGGAGAAATTCAGAG 159
   |||||
87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGlyG 103
   |||||
160 AATATAGGCGGAGACCTGTCATATTTTGCACACGGTTTGCCTCCATC 209
   |||||
103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
   |||||
210 AGCCACAACTGGATCTCCAACTGCGCCCAACACAGCCTGGCCTTCATCC 259
   |||||
120 euAlaSpAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
   |||||
260 TGGCCGACCGGGGTACGACGTGTGCTGGGAAACAGACAGGGGCAACACC 309
   |||||
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
   |||||
310 TGGCCGAGGAGGAATCTGTACTCTCCGCGACTCCGTCGAATCTGGGC 359
   |||||
153 aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsn 170
   |||||
360 TTTCACCTTTGAGGAGATGCGTAATATGACCTTCGCGCCACCATTGACT 409
   |||||
170 heIleLeuGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSer 186
   |||||
410 TCATCTTGAAGAAACGGGACAGGACAAGCTACACTACGTTGGCCATTCC 459
   |||||
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
   |||||
460 CAGGGCACCCACCATGGTTCTATCGCTTTTCCACCAATCCCAAGCTGGC 509
   |||||
203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
   |||||
510 GAACCGGATCAAAACCTTCTATGTCATAGCTCCCGTTGGCCAGCTGAAGT 559
   |||||
220 isAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetIle 236
   |||||
560 ACACCGAAACCCGTGTTAAACAACTCATGCTGCTCCCTTCGTTCCCTTC 609
   |||||
237 LysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArgPheLeuAr 253
   |||||
610 AAGCTTATATTGGAAACAAATATTC...TACCCACACCACTCTCTTGA 656
   |||||
253 gGln...LeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
   |||||
657 TCAATTTCTCGCCACCGAGGATAGCTCCCGGAGAGCGTGGATCTCTCT 706
   |||||
269 ysSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsn 285
   |||||
707 GCAGCAACGCCCTGTGTTATCATTTGTGGATTTGACACTATGAACCTTGAAC 756
   |||||
286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVa 302
   |||||
757 ATGAGTCGCTTGGATGTGATCTGTACATATCCAGCAGGAACATCGGT 806
   |||||
302 iGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgA 319
   |||||
```

```
807 TCAGAACGTGCTCCACTGCTCCAGGCTGTTAAGTCTGGGAAGTTCCAAG 856
319 laPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
   |||||
857 CTTTGTACTGGGAGAGCCAGCTTCAGAACATGATGCACTATCATCAGAGC 906
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
   |||||
907 ATGCCTCCCTACTACAACCTGCAGACATGCATGTGCCAATCCGAGTGTG 956
352 pThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
   |||||
957 GAACGCTGGCAAGGACTGTGCGCCAGCCCTCAGATGTTGACCTTTTGC 1006
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
   |||||
1007 TTTCAGGCTCCCAATCTCATTTACACAGGAAGATTCCTCTTACAAT 1056
386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnG 402
   |||||
1057 CACTTGGACTTATCTGGCCATGGATGCCCTCAAGCGGTTTACATGA 1106
402 uIleIleHisLeuMet 407
   |||||
1107 AATTGTTTCCATGATG 1122
```

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:US-09-356-952-12

seq_documentation_block:

; Sequence 12, Application US/09356952

; Patent No. 6117663

; GENERAL INFORMATION:

; APPLICANT: Horlack-Sjodin, Ann

; APPLICANT: Margarit, S. M.

; APPLICANT: Bor-Sogli, Dafna

; APPLICANT: Cole, Philip

; APPLICANT: Kuriyan, John

; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 600-1-228N

; CURRENT APPLICATION NUMBER: US/09/356,952

; CURRENT FILING DATE: 1999-07-19

; EARLIER APPLICATION NUMBER: 60/093,631

; EARLIER FILING DATE: 1998-07-21

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 12

; LENGTH: 43676

; TYPE: DNA

; ORGANISM: Saccharomyces cerevisiae

; US-09-356-952-12

alignment_scores:

Quality: 271.50 Length: 376

Ratio: 1.263 Gaps: 16

Percent Similarity: 57.181 Percent Identity: 25.532

alignment_block:

US-09-333-159-47 x US-09-356-952-12

Align seg 1/1 to: US-09-356-952-12 from: 1 to: 43676

```
61 GluGluTyrGluValAlaThrGluAspGlyTyrIleLeuSerVal...As 76
   |||||
14769 GAGGAATTTAGATTGGAGACCGGAGATGGGTTTGTATTAGATTTGTGCA 14818
76 nArgIleProArgGlyLeuValGlnProLysLysThrGlySerArgProV 93
   |||||
14819 CTTGATACCAAAATATAGAACGACAGATCTTGACAAAGAAGAGGCCAC 14868
93 alValLeuLeuGlnHisGlyLeuValGlyAlaSerAsnTrpIleSer 109
   |||||
```

```
14869 CCATTTTGATGCTACATGGCGCTTTTGCAGAACGACGTGCTGCTTCCGATCC 14918
110 AsnLeuProAsnAsnSerLeuGlyPheLeuLeuAlaAspAlaGlyPheAs 126
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
14919 AAT...GGTAGAAAATCTCTGGCATATTCTTGATCAATCCGGTTACGA 14965
126 pValTrpMetGlyAsnSerArgGlyAsn.....AlaTrpSerArg. 139
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
14966 CATATGGTTAGGAATACAGATCGCGGTTTAGCCCGAATGGGAACGAAG 15015
140 ..LysHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15016 CGAAGTACCGACACTAGCT.....TCCAGTGGGACTGGGAC 15053
156 TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15054 CTTCGCGAGATGGTAAAGTAGACTCTGACCCCTTTTGATGATACCGTGT 15103
172 uGlnLysThrGlyGlnGluLysIleTyrValGlyTyrSerGlnGlyT 189
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15104 AGCTAAGACGAGTTGAAAGCTTACTTTGATCTCGCAATCTCAGGGCA 15153
189 hrThrMetGlyPheIleAlaPhe..... 196
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15154 CTACACAGGGGTTTATGGCTGTGCTCAACGAAGATAGTTTTCCTCC 15203
197 .....SerThrMetProGluLeuAlaGlnLysIleLysMetTyrPh 210
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15204 GGTTCCGGGATCTAAGAATCTTTTTCACCTTCAAGATCGCAACTATAT 15253
210 eAlaLeuAlaProIleAlaThrValLysHisAlaLysSerProGlyThrL 227
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15254 TGCCTTGGCCCC.....GCAGTGTATCTCTGCTGCC. 15284
227 yspHeLeuLeuLeuProAspMetMetIleLys..... 237
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15285 .....TTACTTAACGAGAAATGTTGTAAAGCTTATGACAAAGAAATC 15329
238 .....GlyLeuPheGlyLysGluPheLeuTyrGlnThrArgPh 251
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15330 GAAATCCCTGGTCTTGTGTGAACAGAGCTTTTTCAGATAATGATGAT 15379
251 eLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnI 268
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15380 TGTAAAGAACTGTGCGTT.....GTGAGAGCTTGTCTCCTTGG 15420
268 leCysSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMet 284
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15421 TTTGTTACACCATCTTCAATATACCTGTTGATTGGAACGATACCTTTGG 15470
285 AsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSe 301
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15471 GATACCGCATTAAGAGATCGCCATTTCTGTTTTCGCCAGTCCCATGTTTC 15520
301 rValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuA 318
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15521 AGTGAAG...TTGATGAATGGTGGCTGTACCCGACCCCAACAGGTA. 15566
318 rGAlaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGln 334
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15567 ..AGTTTAAATTTGGTTCCCATAGATG.....TTC 15596
335 ProThrProValArgTyrArgValArgAspMetThrValProThrAla.. 350
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15597 CCCGACATGTAAAGTGGTTTTCAGACGCATCAAAAGCCCAATATCTA 15646
351 .MetTrpThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysM 367
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15647 CTGTTTGTTCCAAGCAAGATAGATTG.....GTGACCGGAGAA 15687
367 etLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsn..... 380
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15688 GACTAATCAATCAATTCTCGTCAATGCGAGTCGATGTCACACTACAAGATC 15737
```

```
381 .....lleProGluTrpAlaHisValAspPheIleTrpGlyLeuAspAl 395
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15738 TGGTACATTGATGATGATGCCATATTGATGCTCTATGGGCACATGATGT 15787
395 aProHisArgMetTyrAsnGluIleIle 404
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15788 CATAGAGAGAATTGGTAACCAATTTTA. 15815
seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-751-782-2
seq_documentation_block:
; Sequence 2, Application US/08751782
; Patent No. 581352
; GENERAL INFORMATION:
; APPLICANT: Heintz, Nathaniel
; APPLICANT: Gubbay, Johnathan
; APPLICANT: Skinner, Michael
; TITLE OF INVENTION: A cDNA Library Prepared during
; TITLE OF INVENTION: Regression of Rat Prostate and Methods of Use Thereof
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,782
; FILING DATE: 18-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson, Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; IMMEDIATE SOURCE:
; CLONE: 10.2
US-08-751-782-2
```

```
alignment_scores:
Quality: 155.00 Length: 50
Ratio: 3.780 Gaps: 0
Percent Similarity: 82.000 Percent Identity: 56.000
alignment_block:
US-09-333-159-47 x US-08-751-782-2 ...
Align seg 1/1 to: US-08-751-782-2 from: 1 to: 178
139 ArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSe 155
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7 AGGAAACATGTGCGCCTAGACCCAGATTCTAAAGAAATTTGGGATTTAG 56
155 rTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleL 172
```



```
101 lGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsn..... 114
117 GAATGGAAGAGAACTCTTCTAGAAATTTGCAATCAATTTAGAGAAG 166
115 .....SerLeuGlyPhe.....IleLeuAlaAspAla 123
167 ATTTGGACAACTGGCTTCGAGATGAACACTGTATCAATGGCTCAATCA 216
124 GlyPheAspValTrp.....MetGlyAsnSer.....ArgG1 134
217 GCAATCGAAATTAATCTTATCGAAGGAAATTCGGCGTCAGAACGAATGAA 266
134 yAsnAlaTrpSerArgLysHisLys.....ThrLeus 145
267 TAATGCGAAATCCCGTCGCGAAGACCTGCTGGAAATGGAAATATCAATG 316
145 erileAspGlnAspGluPheTrpAlaPheSerTyAspGluMetAlaArg 161
317 AAATGAGGAAGACGATCAGATCAATGCTATCATCTGATTCACAAAA 366
162 PheAspLeuProAlaValIleAsnPheIleLeuGlnLysThrGlyGlnG1 178
367 TGTGATCTCCCTGACCTCGAATATATATACCAAAAAAGATAACACGGA 416
178 uLysIleTyTrpValGlyTySerGlnGlyThrThrMetGlyPheIleA 195
417 ATCGAAAAACGATGGAGCATCTCGGGATAATGATTCGGA..... 459
195 laPheSerThrMetProGluLeuAlaGlnLysIleLysMetTyTrpPheAla 211
460 .....AAATGGAAGCTTTACTCTAGCA 480
212 LeuAlaProIleAlaThrValLysHisAlaLysSerProGlyThrLysph 228
481 GCTGGA..... 486
228 eLeuLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysLysGluP 245
487 .....TTTATATCTTTTTCACGGTGTGCGTTCGAAGCGTGATG 526
245 heLeuTyTrpThrArgPheLeuArgGlnLeuValIleTyTrpLeuCysGly 261
527 TTCCTC.....ACGAATTTGAGATGAGCTATCCGATTTATACATATATG 570
262 GlnValIleLeuAspGlnIleCysSerAsnIleMetLeuLeuLeuGlyG1 278
571 AGAGTGGATGCACGGAAGATGGCTCAATGTAAAGTTCTCTTGGAGC 620
278 yPheAsnThrAsn.....AsnMetAsnMetSerArgAlaSerValT 292
621 TATCAATGAGAATATGAAGCTGAATTTGTAATGTGAAGAGAGGC..... 663
292 yAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTrp 308
664 .....CAATCTACGATTAGTGG 681
309 SerGlnAlaVal.....AsnSerGlyGlyLeu..... 317
682 GCTGATCTATTCGAGAAAAATGAATAGCCACACAGTTGATCTTATCAT 731
318 .....ArgAlaPheAspTrpGlySerGlu 325
732 TCATAATATTGAAGCTCTGATGGAGAGTGTAT 765
```

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq.us-08-484-106-21

seq_documentation_block:
; Sequence 21, Application US/08484106
; Patent No. 5614618
; GENERAL INFORMATION:
; APPLICANT: STILLMAN, Bruce
; APPLICANT: BELL, Stephen P

```
APPLICANT: KOBAYASHI, Ryuji
APPLICANT: RINE, Jasper
APPLICANT: FOSS, Margit
APPLICANT: McNALLY, Francis J
APPLICANT: LAURENSEN, Patricia
APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J
APPLICANT: GAVIN, Kimberly
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSER: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
City: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,106
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oshman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 1676 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 13..1302
US-08-484-106-21
```

alignment_scores:

Quality: 105.00 Length: 278
Ratio: 0.802 Gaps: 13
Percent Similarity: 47.122 Percent Identity: 21.942

alignment_block:

US-09-333-159-47 x US-08-484-106-21 ..

Align seg 1/1 to: US-08-484-106-21 from: 1 to: 1676

```
85 ProLysLysThrGlySerArgProValValLeuLeuGlnHisGlyLeuVa 101
||||| |||||||
85 CCAGAAAAGAGAGATCCAGA.....CAGAAAAGAC 116
101 lGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsn..... 114
||||| |||||||
117 GAATGGAAGAGAACTCTTCTAGAAATTTGCAATCAATTTAGAGAAG 166
115 .....SerLeuGlyPhe.....IleLeuAlaAspAla 123
||||| |||||||
167 ATTTGGACAACTGGCTTCGAGATGAACACTGTATCAATGGCTCAATCA 216
124 GlyPheAspValTrp.....MetGlyAsnSer.....ArgG1 134
||||| |||||||
217 GCAATCGAAATTAATCTTATCGAAGGAAATTCGGCGTCAGAACGAATGAA 266
```

```
134 yAsnAlaTrpSerArgLysHisLys.....ThrLeus 145
:||||| :||||| :|||
267 TAATCGAAATCCCGTCTCGAAGACGTGCTGGAATGAATACTAGAG 316
:||||| :||||| :|||
145 erIleAspGlnAspGluPheTrpAlaPheSerTyAspGluMetAlaArg 161
:||||| :||||| :|||
317 AAATGAGGAAGACGATGAGATCATGATATGCTATCATCTGATTTACAAAA 366
:||||| :||||| :|||
162 PheAspLeuProAlaValIleAsnPheIleLeuGlnLysThrGlyGlnG 178
:||||| :||||| :|||
367 TGTGATCTCCCTGGACTTCGGAATATATATACCAAAAGATACACCGGA 416
:||||| :||||| :|||
178 uLysIleTyTrpValGlyTrpSerGlnGlyThrThrMetGlyPheIleA 195
:||||| :||||| :|||
417 ATTCGAAAACGATGGAGCATCTCGCGGATATGATTCGGA..... 459
:||||| :||||| :|||
195 laPheSerThrMetProGluLeuAlaGlnLysIleLysMetTyTrpPheAla 211
:||||| :||||| :|||
460 .....TTATATATCTTTTCACGGTCTCGGTTCTGAAGCGTGATG 480
:||||| :||||| :|||
212 LeuAlaProIleAlaThrValLysHisAlaLysSerProGlyThrLysPh 228
:||||| :||||| :|||
481 GCTGGA..... 486
:||||| :||||| :|||
228 eLeuLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysLysGluP 245
:||||| :||||| :|||
487 .....TTATATATCTTTTCACGGTCTCGGTTCTGAAGCGTGATG 526
:||||| :||||| :|||
245 heLeuTyGlnThrArgPheLeuArgGlnLeuValIleTyTrpLysCysGly 261
:||||| :||||| :|||
527 TTCTC.....ACAGATTTGAGATGAGCTATCGGATTATACATATATG 570
:||||| :||||| :|||
262 GlnValIleLeuAspGlnIleCysSerAsnIleMetLeuLeuGlyG 278
:||||| :||||| :|||
571 AGAGTGGATGACGCGAAAGATGGGCTCARTGTAAGTCTCTCTGGAGC 620
:||||| :||||| :|||
278 yPheAsnThrAsn.....AsnMetAsnMetSerArgAlaSerValT 292
:||||| :||||| :|||
621 TATCAATGAGATATGAGCTGAATGTAATGTGAAGAGAGCC..... 663
:||||| :||||| :|||
292 yAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTrp 308
:||||| :||||| :|||
664 .....CAATCTACGATTAGTTGG 681
:||||| :||||| :|||
309 SerGlnAlaVal.....AsnSerGlyGluLeu..... 317
:||||| :||||| :|||
682 GCTGATCTATTCCGCAAAAAATGAATAGCAACCAAGTTGATTCTTATCAT 731
:||||| :||||| :|||
318 .....ArgAlaPheAspTrpGlySerGlu 325
:||||| :||||| :|||
732 TCATAATATTGAGCTCCTGATTGGAGAGTGAT 765
:||||| :||||| :|||

seq_name: /cgn2_6/ptodata/1/ins/5B_COMB.seq:US-08-602-359A-24

seq_documentation_block:
; Sequence 24, Application US/08602359A
; Patent No. 5942430
; GENERAL INFORMATION:
; APPLICANT: ROBERTSON, Daniel E.
; APPLICANT: MURPHY, Dennis
; APPLICANT: REID, John
; APPLICANT: MAFFIA, Anthony
; APPLICANT: LINK, Steven
; APPLICANT: SWANSON, Ronald V.
; APPLICANT: WARREN, Patrick V.
; APPLICANT: KOSMOTKA, Anna
; TITLE OF INVENTION: ESTERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 EXECUTIVE SQUARE, STE 1400
; CITY: LA JOLLA
; STATE: CALIFORNIA
```

```
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,359A
FILING DATE: February 16, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HAILE, LISA A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/010001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1041 NUCLEOTIDES
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: GENOMIC DNA
US-08-602-359A-24

alignment_scores:
Quality: 102.00 Length: 180
Ratio: 1.052 Gaps: 8
Percent Similarity: 53.889 Percent Identity: 22.778

alignment_block:
US-09-333-159-47 x US-08-602-359A-24 ..
Align seg 1/1 to: US-08-602-359A-24 from: 1 to: 1041

20 LeuIleLeuValAlaLysMetPheGlnArgAsnValAsnSerValHisMe 36
||||| :||||| :|||
178 CTGCTCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 227
36 tProThr.....LysAlaValAspProGluAlaPheMetAsnI 49
:||||| :||||| :|||
228 CCGCGCTTATGTGGCTATAGTAGTCTGCACCGGAGAGCTGTCGCCCG 277
49 leSerGluIleIleGlnHisGlnGlyTyProCysGluGlnTyGluVal 65
:||||| :||||| :|||
278 TTGCCCGC.....GAGGGGCTCACCTACAAAGAGTTCACGGTG 315
66 AlaThrGluAspGlyTyTrpIleLeuSerValAsnArgIleProArgGlyLe 82
:||||| :||||| :|||
316 ACCGCGGAGGATGCTTGTGTT.....CGGGCTG 347
82 uValGlnProLysLysThrGlySerArgProValValLeuLeuGlnHisG 99
||| :||||| :|||||
348 GGTGCTGGGCCCGCGCTGGGGGCAACCGGCTGCTGCTGCTGCTGCTGCTG 397
99 lyLeuValGly...GlyAlaSerAsnTrpIleSerAsnLeuProAsn 114
|| :||||| :|||||
398 GGTATCTGCTGGCTCGCGCTCGCGCCCTACATGCTGCTGCTGCTGCTGCTG 447
115 SerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyAs 131
||||| :||||| :|||
448 .....CTCGTGGAGTGGGGGTACCCCGTGGTGTGTTGTTGTA 482
131 nSerArgGlyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAspG 148
:||||| :||||| :|||
483 CTTCCGGGCCACGCGGAGAGCGGGGCTCGACGAGGATTGGG..... 525
```

148 InAspGluPheTrpAlaPheSerTyrAspGluMetAlaAlaArgPheAspLeu 164
 526CCCCGGGAGGTGCTGGATGCC 546
 165 ProAlaValIleAspPheIleLeuGlnLysThrGlyGlnGluLysIleTyr 181
 547 CGGGCTGTGGGGGTATGCTCGGAGCGGTTCCTCCGCCGCCGGGATAAT 596
 181 rTyrValGlyTyrSerGlnGlyThrThrMetGlyPheIle 194
 597 ATTGGGGGTTCACTGGCGGCGCTGTAGCGATCGT 636

seq_name: /cqn2_6/ptodata/1/ina/5A_COMB.seq:US-08-232-519-1

seq_documentation_block:

```

; Sequence 1, Application US/08232519
; Patent No. 5484725
; GENERAL INFORMATION:
; APPLICANT: Kageyama, Bunji
; APPLICANT: Nakae, Masanori
; APPLICANT: Yagi, Shigeo
; TITLE OF INVENTION: NORBORANE TYPE ESTER HYDROLASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,519
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-96286
; FILING DATE: 22-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: SHGN-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1167 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Acetobacter pasteurianus
; STRAIN: ATCC 12873
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1164
; US-08-232-519-1

```

alignment_scores: Quality: 98.50 Length: 400
 Ratio: 0.518 Gaps: 20
 Percent Similarity: 47.500 Percent Identity: 20.250

alignment_block:
 US-09-333-159-47 x US-08-232-519-1

Align seg 1/1 to: US-08-232-519-1 from: 1 to: 1167

```

38 ThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleIleG1 54
   ||| ||||| ||||| ||||| |||||
250 ACGCAGATGTGGATATTGAAGCCCTTATC..... 279
54 nHisGlnGlyTyrProCysGluGluTyrGlu.....ValAlaThrGluA 69
   ||||| ||||| ||||| ||||| |||||
280 ....AAAACTTTCATGCAGAAACCCACAGATGCGCTGCAACAGAAG 325
69 spGlyTyrIleLeuSerValAsnArgIleProArgGly.....LeuVal 83
   ||||| ||||| ||||| ||||| |||||
326 ATGCTCTGCCGGGGAACCCAGAGGTTACGGTAGCGGAACACACAGCTA 375
84 GlnProLysLysThrGlySerArg.....ProValValLeuLeuG1 97
   ||||| ||||| ||||| ||||| |||||
376 AATGTGCGTGATGTGGCAGCAGGAGGCGCCCATTTGCTGGTG... 423
97 nHisGlyLeuValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnA 114
   ||||| ||||| ||||| ||||| |||||
424 .CACGGITTTGGCGGAGATATCAGCACTGGCTG.....CTCACACAGG 466
114 snSerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGly 130
   ||||| ||||| ||||| ||||| |||||
467 ATGCTTGGCGGCAGAAAGCGCGTAATTGGCTTTGATCTG..... 507
131 AsnSerArgGlyAsnAlaTrpSerArgLysHis.....LysThrLe 144
   ||||| ||||| ||||| ||||| |||||
508 .....SCGGGGCATGGGGCTTCCTCTAAACACGTGGGCACAGCACGCT 551
144 uSerIleAspGlnAspGluPheTrpAlaPheSerTyrAspGluMetAla 161
   ||||| ||||| ||||| ||||| |||||
552 GCGCTT..... 558
161 rgPheAspLeuProAlaValIleAsnPheIleLeuGlnLysThrGlyGln 177
   ||||| ||||| ||||| ||||| |||||
559 .....TTGGCGGTGTGGTAAGCGAATTGCTGCAAAACCTTAAATA 600
178 GluLysIleTyrValGlyTyrSerGlnGlyThrThrMetGlyPheI1 194
   ||||| ||||| ||||| ||||| |||||
601 GAAAAAGCCCATGTGGTGGCCATCTTTGGGG.....GGCGGCAT 641
194 eAlaPheSerThrMetProGluLeuAlaGlnLysIleLysMetTyrPheA 211
   ||||| ||||| ||||| ||||| |||||
642 TGCCCTGACCTGCTGCGAGATCACCTGATCAGGTTGCCAGCCTGAACC 691
211 laLeuAlaProIleAlaThrValLysHisAlaLysSerProGlyThrLys 227
   ||||| ||||| ||||| ||||| |||||
692 TTTTGGCCCCAGCGCGGTG.....GGAAGGATGTGAATGCAGAT 732
228 PheLeuLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysLysG1 244
   ||||| ||||| ||||| ||||| |||||
733 TTTATC.....AGCCCATTTGTGGATAGTA 758
244 uPheLeuTyrGlnThrArgPheLeuArgGlnLeuValIleTyr..... 258
   ||||| ||||| ||||| ||||| |||||
759 AAGCAGCGCGCATATGAAGGCTGTTTCAATGCTGCTGATATAACAAG 808
259 ..LeuCysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMetLeu 274
   ||||| ||||| ||||| ||||| |||||
809 CCCTAGTGGCGCCGAAGATGGTGGATGCCGTG..... 840
275 LeuLeuGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlaSerVa 291
   ||||| ||||| ||||| ||||| |||||
841 .....CTGCGTGCACGTAGGCTAGATGGCGCGCGGGAT..... 873
291 lTyrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHist 308
   ||||| ||||| ||||| ||||| |||||
874 ....GCCCTGCACGTATTGCTAAAGCGTGTTCCTCCCAAC..... 909
308 rpSerGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlySer 324
   ||||| ||||| ||||| ||||| |||||

```

910GGGCATCAGCGGATATCTGCACCTCGGTGCTAGCTGGGGCG 951
325 GluThrLysAsnLeuGluLysCysAsnGlnProThrProValArgTyrAr 341
|||||
952 GAACA..... 957
341 gValArgAspMetThrValProThrAlaMetTrpThrGlyGlyGlnAspT 358
||||| : : : : :
958CCTACCCAGATTTCTGGGGCAAGGAAGATG 988
358 rpLeuSerAsnProGluAspValLysMetLeuLeuSerGluValThrAsn 374
: : : : :
989 AAATCTTTCTCTCCAAACCGCGCTGCCCTGCCAGATGTCATCCCCGTG 1038
375 LeuIleTyrHisLys.....AsnIleProGluTrpAlaHisValAspPh 389
: : : : :
1039 ACAGTGTATGAAGAAACACAGGCCATCTGCCGCGCTTGAACATGCAACAGA 1088
389 eIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGluIleHisL 406
: : :
1089 TGTG.....AACAAAGCCATTGCC 1108
406 eu...MetGlnGlnGluThrAsnLeuSerGlnGlyArgCysGluAla 421
|| : : : : :
1109 TGTGTAAGAGACCCGAGCGCGCTGAGCATGGCCCGGATGGACCG 1158

OM of: US-09-333-159-47 to: EST:* out_format : pfs

Date: Sep 1, 2001 10:13 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlip
-DB=CG2_1/USPTO.spool/US09333159/runat_310822001_122330_1373/app_query.fasta_1.488
-QF=FASTA-FASTA-SUFFIX=p2n.rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000
-GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEX=7.000 -START=1 -MATRIX=blos62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=200000000 -USER=US09333159 -CGN1_1_3139 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-333-159-47

Query length: 423

Database: EST:*

Database sequences: 10238115

Database length: 431459454

Search time (sec): 1357.660000

score_list:	Strd Orig	ZScore	EScore	Len	Documentation
gb_htc:AK019504	+ 1836.00	3851.43	2.1e-205	2927	AK019504 Mus musculus 0 day ne
gb_htc:AK009571	+ 1139.50	2383.04	1.3e-123	1349	AK009571 Mus musculus adult ma
gb_htc:AK009875	+ 1139.50	2383.04	1.3e-123	1349	AK009875 Mus musculus adult ma
gb_htc:AK010093	+ 1139.50	2383.03	1.3e-123	1350	AK010093 Mus musculus adult ma
gb_htc:AK010139	+ 1139.50	2383.03	1.3e-123	1350	AK010139 Mus musculus adult ma
gb_htc:AK009431	+ 1139.50	2383.01	1.3e-123	1352	AK009431 Mus musculus adult ma
gb_htc:AK009479	+ 1138.50	2380.97	1.7e-123	1343	AK009479 Mus musculus adult ma
gb_htc:AK009473	+ 1138.50	2380.95	1.7e-123	1345	AK009473 Mus musculus adult ma
gb_htc:AK009523	+ 1138.50	2380.94	1.7e-123	1346	AK009523 Mus musculus adult ma
gb_htc:AK009729	+ 1138.50	2380.94	1.7e-123	1346	AK009729 Mus musculus adult ma
gb_htc:AK010116	+ 1138.50	2380.93	1.7e-123	1347	AK010116 Mus musculus adult ma
gb_htc:AK010058	+ 1138.50	2380.92	1.7e-123	1348	AK010058 Mus musculus adult ma
gb_htc:AK009300	+ 1138.50	2380.91	1.7e-123	1349	AK009300 Mus musculus adult ma
gb_htc:AK009474	+ 1138.50	2380.91	1.7e-123	1349	AK009474 Mus musculus adult ma
gb_htc:AK009546	+ 1138.50	2380.91	1.7e-123	1349	AK009546 Mus musculus adult ma
gb_htc:AK009573	+ 1138.50	2380.91	1.7e-123	1349	AK009573 Mus musculus adult ma
gb_htc:AK009459	+ 1138.50	2380.91	1.7e-123	1350	AK009459 Mus musculus adult ma
gb_htc:AK009525	+ 1138.50	2380.91	1.7e-123	1350	AK009525 Mus musculus adult ma
gb_htc:AK009773	+ 1138.50	2380.91	1.7e-123	1350	AK009773 Mus musculus adult ma
gb_htc:AK010061	+ 1138.50	2380.91	1.7e-123	1350	AK010061 Mus musculus adult ma
gb_htc:AK010125	+ 1138.50	2380.91	1.7e-123	1350	AK010125 Mus musculus adult ma
gb_htc:AK009428	+ 1138.50	2380.90	1.7e-123	1351	AK009428 Mus musculus adult ma
gb_htc:AK010019	+ 1138.50	2380.85	1.7e-123	1356	AK010019 Mus musculus adult ma
gb_htc:AK010035	+ 1138.50	2380.85	1.7e-123	1356	AK010035 Mus musculus adult ma
gb_htc:AK010203	+ 1136.50	2376.66	2.9e-123	1350	AK010203 Mus musculus adult ma
gb_htc:AK010124	+ 1135.50	2374.63	3.8e-123	1340	AK010124 Mus musculus adult ma
gb_htc:AK009413	+ 1135.50	2374.59	3.8e-123	1345	AK009413 Mus musculus adult ma
gb_htc:AK010236	+ 1135.50	2374.57	3.8e-123	1347	AK010236 Mus musculus adult ma
gb_htc:AK009937	+ 1133.50	2372.43	5.0e-123	1349	AK009937 Mus musculus adult ma
gb_htc:AK009537	+ 1133.50	2370.30	6.6e-123	1350	AK009537 Mus musculus adult ma
gb_htc:AK009359	+ 1132.50	2368.21	8.1e-123	1346	AK009359 Mus musculus adult ma
gb_htc:AK009544	+ 1131.50	2366.10	1.1e-122	1345	AK009544 Mus musculus adult ma
gb_htc:AK010103	+ 1130.50	2363.96	1.5e-122	1347	AK010103 Mus musculus adult ma
gb_htc:AK009560	+ 1129.50	2361.85	1.9e-122	1346	AK009560 Mus musculus adult ma
gb_htc:AK010026	+ 1129.50	2361.82	2.0e-122	1349	AK010026 Mus musculus adult ma
gb_htc:AK010231	+ 1128.50	2359.71	2.6e-122	1348	AK010231 Mus musculus adult ma
gb_htc:AK009437	+ 1127.50	2357.60	3.4e-122	1346	AK009437 Mus musculus adult ma
gb_htc:AK010036	+ 1125.50	2353.34	5.8e-122	1348	AK010036 Mus musculus adult ma
gb_htc:AK010148	+ 1124.50	2351.21	7.6e-122	1349	AK010148 Mus musculus adult ma
gb_htc:AK010106	+ 1121.50	2344.85	1.7e-121	1349	AK010106 Mus musculus adult ma
gb_htc:AK010173	+ 1120.50	2342.75	2.3e-121	1346	AK010173 Mus musculus adult ma

gb_htc:AK010123 + 1110.50 2321.50 3.4e-120 1350 AK010123 Mus musculus adult
gb_htc:AK009923 + 1109.50 2318.87 4.8e-120 1407 AK009923 Mus musculus adult
gb_htc:AK010091 + 1086.50 2270.60 2.4e-117 1347 AK010091 Mus musculus adult
gb_htc:AK010110 + 1075.50 2247.21 4.7e-116 1353 AK010110 Mus musculus adult
seq_name: gb_htc:AK019504

seq_documentation_block: 2927 bp mRNA HTC 08-FEB-2001
LOCUS AK019504
DEFINITION Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched
library, clone:4632427C23, full insert sequence.

ACCESSION AK019504
VERSION 1 GI:12859754
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) 0 day neonate skin cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:4632427C23.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (sites)
REFERENCE Carninci,P. and Hayashizaki,Y.
AUTHORS High-efficiency full-length cDNA cloning
JOURNAL Methods Enzymol. 303, 19-44 (1999)
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374

REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Kikuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T.,
Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T.,
Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Iizawa,M., Onara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913

REFERENCE 4 (sites)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and
FAJOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 2927)

Akachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
Imotani,K., Ishii,Y., Itoh,M., Iizawa,M., Kato,H., Kawai,J.,
Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
Sugabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission

Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia, Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.

with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR.

FEATURES

source

Location/Qualifiers

```
1..1349
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGI:1895763"
/db_xref="MGI:1914967"
/clone="2310032G08"
/sex="male"
/tissue="male"
/tissue="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
52..1239
/notes="putative"
/codon_start=1
/protein_id="BAB26368.1"
/db_xref="GI:1284448"
/translation="MWLLVTSVLSAFGGAHLFGKLPKNEANNVSMQIMTYWGYP
SEEVYEDGYTLGYRIPYGGKNSENIGKRPVAYLQHLIASATNWTNPNNSLA
FILADAGYDVLGNSRGTWSRKNVYSPDSVEFAFSDMAKYDLPATIDFIVQKT
GQKIHVGHSGQTTIGFIATSTNPALAKIKRFYALAPVATVKYTESPKKISLIPK
FLKLVFGKMEPHNYLDQFLGTEVCSRELLDLCSNALFIECGEDKKNLNVRFDV
YLGNFAGSTQDLFWAQLAKSGKQAYNWGSPQNMLHYNOKTPPYDYVSAWTVPI
AVWGHGHDILADPDQVAMLLPKLPNLLYHKEILPYNHLDFIWAMDAPQEVYNEIVTM
AED"
```

CDS

```
BASE COUNT 396 a 290 c 284 g 379 t
ORIGIN
```

alignment_scores:

Quality: 1139.50 Length: 411
Ratio: 3.485 Gaps: 4
Percent Similarity: 79.562 Percent Identity: 52.555

alignment_block:

US-09-333-159-47 x AK009571

Align seg 1/1 to: AK009571 from: 1 to: 1349

```
7 ArgGlnTrpIleValSerHisArgMetGluMetTrpLeuLeuLeuVal 23
|||||
34 AGGCACCTGG.....CACTCAAGATGTGGCTGCTATTAGTAAC 71
|||||
23 AlaTyrMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
|||||
72 AGTGTGCTATCTGCTGATTGGAGTGCACATGGCTATTGGAAACTGG 121
|||||
39 yAlaValAspProGluAlaPheMetAsnIleSerGluIleLeuGlnHis 55
|||||
122 GTCCCAAAACCCCTGAAGCAACATGAATGTTAGTCAGATGATACCTAC 171
|||||
56 GlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspGlyTyrI 72
|||||
172 TGGGATATCCAAAGTGAGGAATATGAAGTTGTACTGAAGATGGCTACAT 221
|||||
72 eLeuSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrG 89
|||||
222 TCTGGGGTCTATAGAAATTCCTATGGGAGAGAAATTCAGAAATATCG 271
|||||
89 lYserArgProValValLeuLeuGlnHisGlyLeuValGlyClyAlaSer 105
|||||
272 GCAAGAGACCTGTGCATATTGTCAGCATGGTTGATTGTCATCAGCCACA 321
|||||
106 AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
|||||
322 AACTGGATTACAAATCTGCCAAACACACAGCCCTGGCCCTTCATTTCTAG 371
|||||
122 pAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerA 139
|||||
372 TCGTGGCTATGATGTGGCTGGGACACAGTCGAGGGAATACATGGTCCC 421
```

```
139 rgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
|||||
422 GGAATAATGTATATCTATTCACACAGCTAGTGAATTCGGCTTTCAGC 471
|||||
156 TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
|||||
472 TTTGATTAATGGCTAATATAGACTCTCCACCCACCATAGACTTCATTTG 521
|||||
172 uGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSerGlnGlyT 189
|||||
522 ACAGAAACTGGACAAGAGAATACACTATGTTGGTCACTCTCAGGSCA 571
|||||
189 hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
|||||
572 CCATATCGGTTTATTGGCTTTTCTACCAATCCTGCTCTGGCTAAATAA 621
|||||
206 IleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaL 222
|||||
622 ATCAAGAGGTTTATGATTAAGTCCAGTTCCTACTGTGAAGTATACAGA 671
|||||
222 sSerProGlyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyL 239
|||||
672 AAGTCCCTTTAAAAAGATTTTCACTTATTCCTAAGTTTCTTCTCAAGGTGA 721
|||||
239 euPheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu 252
|||||
722 TATTGTGTAACAAAATGTTTCACTGCCCCACAACTACTTAGATCAATTTCT 771
|||||
253 ArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleCy 269
|||||
772 GGTACGGAAGTG.....TGCTACGGGAGCTGCTAGATCTTCTCTG 812
|||||
269 sSerAsnIleMetLeuLeuLeuGlyPheAsnThrAsnAsnMetAsnM 286
|||||
813 CAGCAACGCTTTATTCATCTTCTGTGGATTGACAAAGAAACAACTTAATG 862
|||||
286 etSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
|||||
863 TGAGTCGCTTTGATGTATCTAGGCATATATCCAGCAGGAACATCTACT 912
|||||
303 GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
|||||
913 CAAGACCTTTCCACTGGGCACAGCTTCTAAATCTGGGAGCTTCAAGC 962
|||||
319 aPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnProT 336
|||||
963 CTATAACTGGGGAAGTCCATACAGAACATGTTACACTACAAATCAGAAA 1012
|||||
336 hrProValArgTyrArgValArgAspMetThrValProThrAlaMetTrp 352
|||||
1013 CGCCTCCCTACTATGATGTGTGAGCCATGACCGTGCCCAATTCAGTGG 1062
|||||
353 ThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLe 369
|||||
1063 AACGGIGCCCATGACATCTGCTGATCCCAAGATGTCGCAATGCTGCT 1112
|||||
369 uSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAlaH 386
|||||
1113 TCCCAAACTCCCAACCTTCTGTACATAAGGAGATTTCTCCCTACAATC 1162
|||||
386 lsValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGlu 402
|||||
1163 ACCTGGACTTCTATCTGGGCGATGATGGCGCTCAAGAGGTTTACAATGAG 1212
|||||
403 IleIleHisLeuMetGlnGlnGluThrAsn 413
|||||
1213 ATAGTTACCATGATGGCAGACACTAACAGAAAT 1245
```

seq_name: gb_htc:AK009875

seq_documentation_block:

LOCUS AK009875

DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched

HTC

08-FEB-2001


```

library, clone:2310047G15, full insert sequence.
ACCESSION AK009875
VERSION AK009875.1 GI:12844938
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:2310047G15.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods Enzymol. 303, 19-44 (1999)
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20493374
REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuina,T., Tashiro,H., Itoh,M.,
Kikuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T.,
Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T.,
Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
REFERENCE 4 (sites)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 683-690 (2001)
REFERENCE 5 (bases 1 to 1349)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
Miyazaki,A., Nishikawa,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan. (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAAGAGCTCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0 and subtraction to
Rot = 25.0. Second strand cDNA was prepared with the primer adapter
of sequence [5'
GAGAGAGATCTCGAGTTAATTAAATTCCTCCCTCCCTCCCTCCCTCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

```

```

Host: SOLR
FEATURES             Location/Qualifiers
     source            1..1349
     organism="Mus musculus"
     strain="C57BL/6J"
     db_xref="taxon:10090"
     db_xref="MGI:1898366"
     db_xref="MGI:1914967"
     clone="2310047G15"
     sex="male"
     tissue_type="tongue"
     clone_lib="RIKEN full-length enriched mouse cDNA library"
     dev_stage="adult"
     53..1240
     note="putative"
     codon_start=1
     protein_id="BAB26556.1"
     db_xref="GI:12844939"
     translation="WLLLVTSVLSAFGGAHGLFGLGPKNPKANNVSNOMITYWGYV
SEYEVWEDGYILGVRIPTYCKKSENIGKRPVAYLQHLIASATNWTNLENNSLA
FILADAGYDVLNGSRNTWKNYISPDSEVFASFDEMAKIDLPATIDFVQKT
GOEKTHYVGSQCTTIGTAFSTNPALAKKIKRFYAPVATVKYTESPPKKLSLIPK
FLIKVTFGKNMFPHNYLDQFLGTEVCSRELDLCSNALFICGDFKKNLNYSRFDV
YLCHNPAGTSTODLPFWAOLAKSKLQAYNWGSPNLMLHYNOKTPPYDVSAATVPI
AVWNGGHDILADPDQDVAMLLPKLPNLLYHKEILPYNHLDFIWMADAPQEVYNEIVTMM
AED"
BASE COUNT          396 a 290 c 284 g 379 t
ORIGIN
alignment_scores
Quality: 1139.50      Length: 411
Ratio: 3.495          Gaps: 4
Percent Similarity: 79.319 Percent Identity: 52.555
alignment_block
US-09-333-159-47 x AK009875
Align seg 1/1 to: AK009875 from: 1 to: 1349
7 ArgGlnTrpIleValSerHisArgMetGluMetTrpLeuLeuLeuVa 23
|||||
35 AGGCATCGG.....CACTACAAGATGTGGCTGCTATTAGTAAC 72
23 lAlaTrpMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
|||||
73 AAGTGTCTCTCTCATTTGGAGGTGCACATGGCTATTGGAAAACTGG 122
39 ySaLaValAspProGluAlaPheMetAsnIleSerGluIleIleGlnHis 55
|||||
123 GTCCCAAAACCCCTGAAGCAACATGAATGTTAGTCAGATGATAACTTAC 172
56 GlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspGlyTyrI 72
|||||
173 TGGGATATCCAACTGAGGAGATATCAAGTGTCTACTGAAGATGGCTACAT 222
72 eLeuSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrG 89
|||||
223 TCTGGGGTCTATAGATTCCTTATGGGAAGAAAAATTTCTGAGAAATCG 272
89 lySerArgProValValLeuLeuGlnHisGlyLeuValGlyGlyAlaSer 105
|||||
273 GCAGAGACCTGTGGCATATTTCGACATGTTTGTTCATCAGCCACCA 322
106 AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
|||||
323 AACTGGATTACAAATCTGCCAAACAACACCTGGCCTTCATCTAGCACA 372
122 pAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerA 139
|||||
373 TGCTGCCTATGATGTGTGGCTGGGAACACAGTCGAGGGAATACATGTCCTC 422
139 rgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155

```

ACCESSION VERSION KEYWORDS SOURCE	AK010093 AK010093.1 CAF trapper. Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA. clone-lib:RIKEN full-length enriched mouse cDNA library clone:2310067K20.
ORGANISM	Mus musculus
REFERENCE AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE JOURNAL REFERENCE AUTHORS	1 (sites) Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
TITLE JOURNAL MEDLINE REFERENCE AUTHORS	2 (sites) Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itch, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 2049374
TITLE JOURNAL MEDLINE REFERENCE AUTHORS	3 (sites) Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, K., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system- and Hayashizaki, Y. sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913
TITLE JOURNAL MEDLINE REFERENCE AUTHORS	4 (sites) The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) 5 (bases 1 to 1350) Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Haragaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawa, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saigo, H., Saigo, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y. Direct Submission
TITLE JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ , Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGGATCCAGGACGCTTTTTTTTTTTTTTNN 3'], cDNA was prepared by using trihaloetho thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATTCGAGTAAATTAATTCCTCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

seq_documentation_block:			
LOCUS	AK010093	1350 bp	HTC 08-FEB-2001
DEFINITION	Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310067K20, full insert sequence.		

```

FEATURES
  source      Location/Qualifiers
1..1350
  /organism="Mus musculus"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /db_xref="MGI:1898392"
  /db_xref="MGI:1914967"
  /clone="2310067K20"
  /sex="male"
  /tissue_type="tongue"
  /clone_lib="RIKEN full-length enriched mouse cDNA library"
53..1240
  /dev_stage="adult"
  /note="putative"
  /codon_start=1
  /protein_id="BAB26697.1"
  /db_xref="GI:12845298"
  /translation="MWLLVTSVLSAFGGAGHLEGLKGPKNPEANNVNSQMITWYGP
SEYEVVTEGDIYLVGVRIYKKNSENIGKRPVAVLOHGLVASATNMTNLPNNSLA
FILADAGYDVWLGNSRGNTWRSKNVYSDSPSEFPAFSEDEMAKYDLPATIDFVOKT
GQKIHVGHSGGTTIGFIASFNPALAKIRKYPALAPVATVKYTESPFKILSLIPK
FLKVIYFGNKMFMHNYLDQFLGTEVCSRELLDLCSNALFICGDKKLNINVSREYV
YLGNPAGTSTODLEHWAOLAKSGKQAYNWGSPQLQMLHYNQKTPPYDYDSAMTPI
AVNGGHDILADPDQDVAMLLPKLPNLLYHKEILPYNHLDFIWAMDAPQEVINEIVTM
AED"
BASE COUNT      396 a 290 c 286 g 378 t
ORIGIN

alignment_scores:
  Quality: 1139.50      Length: 411
  Ratio: 3.495          Gaps: 4
  Percent Similarity: 79.319 Percent Identity: 52.798

alignment_block:
US-09-333-159-47 x AK010093 ..
Align seg 1/1 to: AK010093 from: 1 to: 1350

7  ArgGlnTrpIleValSerHisArgMetGluMetTrpLeuLeuIleLeuVa 23
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
35 AGGCACCTGG.....CACTACAAGATGTGGCTGCTATTAGTAAC 72

23 lalaTyrrMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
73 AAGTGTGTCTATCTGTCATTGGAGGTGCACATGGCCTATTGGAAAACCTGG 122

39 ysAlaValAspProGluAlaPheMetAsnIleSerGluIleleGlnHis 55
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
123 GTCCCAAAACCCCTGAAGCAACATGAATGTTAGTCAGATGATAACTTAC 172

56 GlnGlyTyrrProCysGluGluTyrrGluValAlaThrGluAspGlyTyrrI 72
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
173 TGGGATATCCCAAGTGAGGATATGCACTGTTACTGAAGATGGCTACAT 222

72 eleuSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrG 89
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
223 TCTGGGGTCTATAGAAATTCCTTATGGGAAGAAAATTCCTGAGAAATATCG 272

89 lySerArgProValValLeuLeuGlnHisGlyLeuValGlyClyAlaSer 105
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
273 GCAGAGACCTGTGGCATATTGGAGCATGGTTGGTTGCATCAGCCACA 322

106 AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
323 AACTGGATTACAAATCTGCCAAACACAGCCTGGCCTTCATCTAGCAGA 372

122 palaglyPheaspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSera 139
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
373 TGCTGGCTATGATGTGTGGCTGGGGAACAGTCGAGGAGGAATACATGCTCC 422

139 rGlyHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

423 GGAAAAATGTATATATTCACCAGACTCAGTGAATTCCTGGCCTTCAGC 472
156 TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
473 TTTGATGAATGGCTAAATATGACCTTCAGCCACCACATAGACTTCATTGT 522
172 uGlnLysThrGlyGlnGluLysIleTyrrValGlyTyrrSerGlnGlyT 189
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
523 ACAGAAACTGGACAAGAGAATACATATGTTGGTCACTCTCAGGGCA 572
189 hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
573 CCATATCGGTTTATTCGCTTTTCTACCAATCCTGCTCTGGCTAAAAAA 622
206 IleLysMetTyrrPheAlaLeuAlaProIleAlaThrValLysHisAlaL 222
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
623 ATCAAGAGGTTTATGATAGCTCAGTGTGCTACTGTGAAGTATACAGA 672
222 sSerProGlyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyL 239
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
673 AAGTCCCTTTAAAAAGATTTCATTTATTCCTAAGTTTCTCTCAAGGTGA 722
239 euPheGlyLysLysGluPheLeu.....TyrrGlnThrArgPheLeu 252
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
723 TATTGGTAAACAAATGTTTCATGCCCCCACTACTTAGATCAATTCTT 772
253 ArgGlnLeuValIleTyrrLeuCysGlyGlnValIleLeuAspGlnIleCy 269
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
773 GGTACGGAAGTG.....TGCTCAGGGAGCTGTAGATCTTCTCTCTG 813
269 sSerAsnIleMetLeuLeuLeuGlyPheAsnThrAsnAsnMetAsnM 286
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
814 CAGCAACGCTTTATTCATCTCTGTGGATTTCAGCAAGAAAAACTTAAATG 863
286 etSerArgAlaSerValTyrrAlaAlaHisThrLeuAlaGlyThrSerVal 302
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
864 TGAGTCGCTTTGATGTGTATCTAGGGCATATCCAGCAGGAACATCTACT 913
303 GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
914 CAAGACCTTTCCACTGGGCACAGCTTGCTAAATCTGGGAAGCTTCAAGC 963
319 apheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnProT 336
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
964 CTATAACTGGGGAGTCCATACAGAACATGTTACACTACAATCAGAAAA 1013
336 hrProValArgTyrrArgValArgAspMetThrValProThrAlaMetTp 352
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1014 GCCTTCCTACTATGATGTGTCAGCCATGACCGTCCCAATTCGCAATGCTG 1063
353 ThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLe 369
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1064 AACGGTGGCCATGACATCTGCTGATCCCAAGATGTCGCAATGCTGCT 1113
369 uSerGluValThrAsnLeuIleTyrrHisLysAsnIleProGluTrpAlaH 386
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1114 TCCCAAACTCCCAACCTTCTGTACCATAAAGAGATTTCTCCCTACAAT 1163
386 lsValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrrAsnGlu 402
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1164 ACCTGGACTTCATCTGGGGATGGATGCCCTCAGAGAGGTTTACAATGAG 1213
403 lleIleHisLeuMetGlnGlnGluThrAsn 413
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1214 ATAGTTACCATGATGGCAGAACACTAACAGAAT 1246

```

seq_name: gb_hlc:AK010139

seq_documentation_block:

LOCUS AK010139 1350 bp mRNA HTC

DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched

library, clone:2310069P19, full insert sequence.

ACCESSION AK010139

08-FEB-2001


```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGI:1894047"
/db_xref="MGI:1914967"
/clone="2310021816"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
55.1242
/notes="putative"
/codon_start=1
/protein_id="BAB26283.1"
/db_xref="GI:12844223"
/translation="MWLLVTSVLSAFGAHGLFGKLPKNEANNVVSQMITYWGYP
SEYEVVTDGVTILGYRIPYKKNSEINIGKRPVAYLOHGLIASATNWTINLPNNSLA
FIADAGYDVLGNSRGNTWSRKNVYSPDSVEFAFSDEMAYDLPATIDFIVOKT
GOEKIHVHSGOCTTIGFTAFSTNPALAKIKRFYALAPVATVKYTESPKKISLPK
FLKVFEGKNHFNHLYLDOFLGTEVCSRELLDLCSNALFICGDFDKKLNUNSRIDV
YLGNPAGTSTODLFHWAOLAKSGKLOAYNMGSPQLNMLHYNOKTPPYDYDVSAINTVPI
AVWNGHDLADPDQDVAMLLPKLPNLLYHKEILYPNHLDFIWAMDAPQEVYNEIVTMM
AED"
BASE COUNT      397 a 291 c 285 g 379 t
ORIGIN

```

```

alignment_scores:
  Quality: 1139.50      Length: 411
  Ratio: 3.495          Gaps: 4
  Percent Similarity: 79.319 Percent Identity: 52.555

alignment_block:
US-09-333-159-47 x AK009431 ..
Align seg 1/1 to: AK009431 from: 1 to: 1352

7  ArgGlnTrpIleValSerHisArgMetGluMetTrpLeuLeuValLeuVa 23
   ::::::::::::::::::::
37 AGGCACCTGG.....CACTACCAAGATGTGGCTGCTATTAGTAAC 74
   ::::::::::::::::::::

23 lAlaTyrrMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
   ::::::::::::::::::::
75 AGTGTGCTATCTGCATTGGAGTGGACATGGCCCTATTGGAAAACCTGG 124
   ::::::::::::::::::::

39 yAlaValaspProGluAlaPheMetAsnIleSerGluIleIleGlnHis 55
   ::::::::::::::::::::
125 GTCCCAAAAACCTCGAAGCAACATGAATGTTAGTCAGATGATACTTAC 174
   ::::::::::::::::::::

56 GlnGlyTyrrProCysGluGluTyrrGluValAlaThrGluAspGlyTyrrI 72
   ::::::::::::::::::::
175 TGGGGATATCCAAAGTCAGGAATATGAAGTTGTTACTGAAGATGGCTACAT 224
   ::::::::::::::::::::

72 eLeuSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrG 89
   ::::::::::::::::::::
225 TCTGGGGTCTATAGAAATTCCTATGGGAAGAAAATTCAGAAATATCG 274
   ::::::::::::::::::::

89 lySerArgProValValLeuLeuGlnHisGlyLeuValGlyGlyAlaSer 105
   ::::::::::::::::::::
275 GCAAGAGACCTCTGGCATATTTCAGCATGTTTTCATTGTCATCAGCCACA 324
   ::::::::::::::::::::

106 AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
   ::::::::::::::::::::
325 AACTGGATTACAAATCTGCCAAACACAGCCTGGCCTTCATTCTAGCAGA 374
   ::::::::::::::::::::

122 pAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerA 139
   ::::::::::::::::::::
375 TCGTGGCTATGATGTGGCTGGGAACAGTCGGAGGAATACATGGTCCC 424
   ::::::::::::::::::::

139 rGlySHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
   ::::::::::::::::::::
425 GGAATAATGTATACATATTACACAGACTCAGTTGAATTTCTGGCCTTCAGC 474
   ::::::::::::::::::::

```

```

156 TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
   ::::::::::::::::::::
475 TTTGATGAATGGCTAAATATGACCTTCCAGCCACCATAGACTTCATTGT 524
   ::::::::::::::::::::

172 uGlnLysThrGlyGlnGluLysIleTyrrValGlyTyrrSerGlnGlyT 189
   ::::::::::::::::::::
525 ACAGAAACCTGGACAGAGAATACACTATGTTGGTCACCTCTCAGGCA 574
   ::::::::::::::::::::

189 hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
   ::::::::::::::::::::
575 CCACATCGGTTTATTATGCTTTTCTACCAATCTGCTGCTGCTAAAAA 624
   ::::::::::::::::::::

206 IleLysMetTyrrPheAlaLeuAlaProIleAlaThrValLysHisAlaL 222
   ::::::::::::::::::::
625 ATCAAGAGGTTTATGCAATAGCTCCAGTTGCTACTGTGAAGTATACAGA 674
   ::::::::::::::::::::

222 sSerProGlyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyL 239
   ::::::::::::::::::::
675 AAGTCCCTTTAAAAAGATTTTCACTTATTCCTAAGTTTCTTCAAGGTGA 724
   ::::::::::::::::::::

239 euPheGlyLysLysGluPheLeu.....TyrrGlnThrArgPheLeu 252
   ::::::::::::::::::::
725 TATTGTGTAACAAAATGTTCAATGCCCCACAACTACTTAGATCAATTCTT 774
   ::::::::::::::::::::

253 ArgGlnLeuValIleTyrrLeuCysGlyGlnValIleLeuAspGlnIleCy 269
   ::::::::::::::::::::
775 GGTACGGAAGTG.....TGCTACGGGAGCTGTAGATCTTCTCTGTG 815
   ::::::::::::::::::::

269 sSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnM 286
   ::::::::::::::::::::
816 CAGCAACGCTTTATTCATCTTCTGTGGATTGACAAAGAAAACCTTAATG 865
   ::::::::::::::::::::

286 etSerArgAlaSerValTyrrAlaAlaHisThrLeuAlaGlyThrSerVal 302
   ::::::::::::::::::::
866 TGAGTCGCTTGTATCTAGGGCATAATCCAGCAGGAACATCTACT 915
   ::::::::::::::::::::

303 GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
   ::::::::::::::::::::
916 CAAGACCTTTCCACTGGGCACAGCTTCTTAATCTGGGAAGCTTCAAGC 965
   ::::::::::::::::::::

319 aPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnProT 336
   ::::::::::::::::::::
966 CTATAACTGGGAAGTCCATACAGAACATGTTACACTACATCAGAAAA 1015
   ::::::::::::::::::::

336 hrProValArgTyrrArgValArgAspMetThrValProThrAlaMetTrp 352
   ::::::::::::::::::::
1016 CGCCTCCCTACTATGATGTGTGACGATGACCGTGCCCAATTCAGTGTGG 1065
   ::::::::::::::::::::

353 ThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLe 369
   ::::::::::::::::::::
1066 AACGGTGGCCATGACATCTGCTGCTGATCCCAAGATGTCGCAATGCTGCT 1115
   ::::::::::::::::::::

369 uSerGluValThrAsnIleTyrrHisLysAsnIleProGluTrpAlaH 386
   ::::::::::::::::::::
1116 TCCCAAACTCCCAACCTTCTGTACCATAGGAGATTCTTCCCTACAATC 1165
   ::::::::::::::::::::

386 lsValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrrAsnGlu 402
   ::::::::::::::::::::
1166 ACCTGGACTTCACTCTGGCGCATGGATGCGCTCAAGAGGTTTACAATGAG 1215
   ::::::::::::::::::::

403 IleIleHisLeuMetGlnGlnGluThrAsn 413
   ::::::::::::::::::::
1216 ATAGTTACCATGATGGCAGAAAGACTTAACAGAAT 1248
   ::::::::::::::::::::

seq_name: gb_htc:AK009479
seq_documentation_block:
LOCUS      AK009479      1343 bp      mRNA      HTC      08-FEB-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
            library, clone:231002K08, full insert sequence.
ACCESSION  AK009479
VERSION    AK009479.1 GI:12844303
KEYWORDS   CAP trapper.

```



```

/db_xref="taxon:10090"
/db_xref="MGD:MGI:189831"
/db_xref="MGD:MGI:1914967"
/clone="2310022H17"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
52 1239
/seq="putative"
/codon_start=1
/protein_id="BAB26312.1"
/db_xref="GI:12844294"
/translation="MALLVTSVLSAFGGAHLGFLGPKPNPEANNVSNOMITYMGYP
SERVEYEDGYILGVRIPIYKGNSENIGKRPVAYLOHGLIASATNWTINLPNNLSA
FILADAGYDNLGNSRGNTWSKNVYSPDSVEFWAFSDENAKYDLPTIDFIVNOKT
GQKIHVGHSGQTIGTAFSTPALAKRIKRFALAPVATVVKYTESPFKLSILPK
FLKVIFGKNMFPHNYLDQFLGTECSRELLDLCSNALFIFCGDDKNNLNSREDF
YLGHPAGTSTQDLFWAQLAKSGKIQAYNWGSPQLQMLHYNOKTPPYDVSAMTYPI
AVWNGGHDILADPDQVAMLLPKLPNLLYHKEILPYNHLDFINAMDAPQEVINEIVTM
AED"
BASE COUNT      395 a 288 c 283 g 379 t
ORIGIN

```

```

alignment_scores:
  Quality: 1138.50      Length: 411
  Ratio: 3.492          Gaps: 4
  Percent Similarity: 79.319 Percent Identity: 52.555

```

```
alignment_block:
```

```
US-09-333-159-47 x AK009473
```

```
Align seg 1/1 to: AK009473 from: 1 to: 1345
```

```

7  ArgGlnTrpIleValSerHisArgMetGluMetTrpLeuLeuLeuVal 23
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
34  AGGCACCTGG.....CACTACAAGATGTGGCTGCTATTAGTAAC 71
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
23  lAlaTrpMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
72  AAGTGTGCTATCGCATTTGGAGGTGCACATGGCCCTATTGGAAAACCTGG 121
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
39  ySAlaValAspProGluAlaPheMetAsnIleSerGluIleIleGlnHis 55
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
122  GTCCCAAAACCTCGAAGCAACATGATGTTAGTCAGATGATACTTAC 171
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
56  GlnGlyTrpProCysGluGluTrpGluValAlaThrGluAspGlyTrpI 72
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
172  TGGGGATATCCAAAGTGAGGAATATGAAGTTGTTACTGAAGATGGCTACAT 221
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
72  eLeuSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrG 89
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
222  TCTGGGGTCTATAGAAATTCCTTATGGGAAGAAAAAATCTGAGAAATCG 271
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
89  lYSerArgProValValLeuGlnHisGlyLeuValGlyAlaSer 105
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
272  GCAAGAGACCTGTGGCATATTGTCAGCATGGTTGATTCATCAGCCACA 321
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
106  AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
322  AACTGATTACAAATCTGCAACACACACGCTGGCCTTCATCTAGCAGA 371
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
122  pAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerA 139
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
372  TGTGCTGCTATGATGTGTGGTGGGAACAGTCGAGGGAATACATGTCCTC 421
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
139  rGlyHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
422  GGAATAATGTATACATTACCCAGACTCAGTTGAATCTGGGCTTTCAGC 471
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
156  TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

472  TTTGATGAATGGCTAAATATATGACCTTCCAGCCACCATAGACTTCATTGT 521
172  uGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSerGlnGlyT 189
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
522  ACAGAAACTGGACAAGAGAATACACTATGTTGGTCACCTCTCAGGCA 571
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
189  hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
572  CCACATCGGTTTATTGCCCTTTCTACCAATCCTGCTCTGGCTAAAAAA 621
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
206  IleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaL 222
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
622  ATCAAGAGGTTTATGCAATAGTCCAGTGTCTACTGTAAGTATACAGA 671
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
222  sSerProGlyThrLysPheLeuLeuProAspMetMetIleLysGlyL 239
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
672  AAGTCCCTTTAAAAGATTTTCACTATTCTTAAGTTTCTTCAAGGTGA 721
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
239  euPheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu 252
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
722  TATTGGTAACAAATGTTTCATGCCCCCACTACTTAGATCAATTCTT 771
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
253  ArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleCy 269
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
772  GGTACCGAAGTG.....TGCTCAGGGGAGCTGCTAGATCTTCTCTG 812
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
269  sSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnM 286
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
813  CAGCAACGCTTATTTCATCTTCTGTGGATTTCACAAGAAAAAATTAATG 862
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
286  etSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
863  TGAGTCGCTTTGATGTATCTAGGGCATAATCCAGCAGGAACATCTACT 912
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
303  GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
913  CAAGACCTTTTCCACTGGGCACAGCTTGCTAAATCTGGAAGCTTCAAGC 962
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
319  apheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 336
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
963  CTATACTGGGAAGTCCATTACAGAACATGTTACACTACAATCAGAAAA 1012
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
336  hrProValArgTyrArgValArgAspMetThrValProThrAlaMetTrp 352
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1013  CGCCTCCCTACTATGATGTGTCAGCATGACCGTGCCAATTCAGGTGG 1062
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
353  ThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLe 369
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1063  AACGGTGGCCATGACATCTGCTGATCCCAAGATGTCGCAATGCTGCT 1112
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
369  userGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAlaH 386
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1113  TCCCAACTCCCAACCTTCTGTACCATTAAGGAGATTTCCCTCAACATC 1162
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
386  isValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGlu 402
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1163  ACTGGACTTCATCTGGGCGATGGATGCCCTCAGAGGTTTACATGAG 1212
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
403  ileIleHisLeuMetGlnGlnGluThrAsn 413
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1213  ATAGTTACCATGATGCAGAAGAACTAACAGAAT 1245
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```
seq_name: gb_hrc:AK009523
```

```
seq_documentation_block:
```

```

LOCUS      AK009523      1346 bp      mRNA      HTC      08-FEB-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
            library, clone:2310028I09, full insert sequence.
ACCESSION  AK009523
VERSION    AK009523.1  GI:12844369
KEYWORDS   CAP trapper.
SOURCE     Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library

```



```

/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
52..1239
CDS
/note="putative"
/codon_start=1
/protein_id="BAB26673.1"
/db_xref="GI:12845239"
/translation="MWLLVTSVLSAFGGAHGLFKLPKPNPEANNMNSQMITTYWGY
SSEYEVVTDGVLGVIRIPYKNSENIGKRPVAYLQHLIASATNWITNLPNNSLA
FILADAGYDGLNSRGNTWSKNNVYSPDSVEFWAFSDENAKYDLDPATIDFVQKT
COEKHYVCHSGGTIGTIGTAEFTNPALAKIKRFYALAPATVKKTESPPKKISLPK
FLKLVIFGNKMFNPYLDOLFTEVCSRELLDLCSNALFIFCGFDRKLNLSREVD
YLGHPAGSTODLFHQAOLSKGLQAYNWGSPLONLMLHYNOKTPPYDYDSAMTVPI
AYWNGCHDLADPOQVAMLLPKLPNLLYHKEILPYNHLDFIWMADAPQEVYNEIVTM
AED"
BASE COUNT      396 a 290 c 284 g 378 t
ORIGIN

```

```

alignment_scores:
  Quality: 1138.50      Length: 411
  Ratio: 3.492          Gaps: 4
  Percent Similarity: 79.319  Percent Identity: 52.555

```

```

alignment_block:
US-09-333-159-47 x AK010058

```

```

Align seg 1/1 to: AK010058 from: 1 to: 1348

```

```

7  ArgGlnTrpIleValSerHisArgMetGluMetTrpLeuLeuVal 23
   |||||
34  AGGCACATGG.....CACTACAAGATGTGGCTATTAGTAAC 71
   |||||
23  lAlaTyrMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
   |||||
72  AAGTGTGCTATCTGCATTTGGAGTGGACATGGCTATTGGGAAACTGG 121
   |||||
39  ySaLaValAspProGluAlaPheMetAsnIleSerGluIleGlnHis 55
   |||||
122  GTCCCAAAACCCCTGAAGCAACATGAATGTTAGTCAGATGATTAAC 171
   |||||
56  GlnGlyTyrProCysGluGlyTyrGluValAlaThrGluAspGlyTyr 72
   |||||
172  TGGGGATATCCCAAGTGAAGATATGAAGTTGTTACTGAAGATGGCTACAT 221
   |||||
72  eLeuSerValAsnArgIleProArgGlyLeuValGlnProLysLysThr 89
   |||||
222  TCTGGGGTCTATAGATTCCTTTATGGGAGAGAAATCTGAGATATCG 271
   |||||
89  lySerArgProValValLeuLeuGlnHisGlyLeuValGlyGlyAlaSer 105
   |||||
272  GCAAGAGACCTGTGGCATATTTCAGCATGTTGTTGATTGTCATGAGCCACA 321
   |||||
106  AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
   |||||
322  AACTGGGATTACAAATCTGCCAAACACAGCCTGGCCTTCATCTAGCAGA 371
   |||||
122  pAlaGlyPheAspValTyrMetGlyAsnSerArgGlyAsnAlaTrpSerA 139
   |||||
372  TGCTGGCTATGATGTGGCTGGGAGACAGTCGAGGGNATACATGGTCCC 421
   |||||
139  rGlySHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
   |||||
422  GGAATAATGTACTATTACCAGACTCAGTTGAATCTCTGGCTTTTCAGC 471
   |||||
156  TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
   |||||
472  TTTGATGAATGGGTAAATATGACCTCCAGCCACCACCATGACTTCATGT 521
   |||||
172  uGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSerGlnGlyT 189
   |||||

```

```

522  ACAGAAACTGGCAACAGAGATACACTATATGTTGGTCACTCTCAGGGCA 571
189  hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
   |||||
572  CCACATCGGTTTATTATGCGCTTTCTACCAATCGCTCTGGCTAAAGAAA 621
206  IleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLys 222
   |||||
622  ATCAAGAGGTTTATGATTTAGCTCCAGTTGCTAGTGAAGATATACAGA 671
222  sSerProGlyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyL 239
   |||||
672  AAGTCCCTTTAAAAAGATTTTCATTTCTTAAGTTCTTCTCAAGGTGA 721
239  euPheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu 252
   |||||
722  TATTTGTAAACAAATGTTTCATGCCCCCAACTACTAGATCAATTCTCT 771
253  ArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleCy 269
   |||||
772  GGTACGGAAGTG.....TGCTACGGGAGCTGTAGATCTTCTCTG 812
269  sSerAsnIleMetLeuLeuLeuGlyPheAsnThrAsnAsnMetAsnM 286
   |||||
813  CAGCAACGCTTTATTCATCTCTGTGGATTGACAAAGAAACCTTAATG 862
286  etSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
   |||||
863  TGAGTCGCTTTGATGCTGATCTAGGCATATATCCAGCAGGACATCTACT 912
303  GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
   |||||
913  CAAGACCTTTTCCACTGGGCACAGCTTCTAAATCTGGGAGCTTCAAGC 962
319  aPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnProT 336
   |||||
963  CTATACTGGGGAAGTCCATTACAGACATGTTACTACTACATCAGAAAA 1012
336  hrProValArgTyrArgValArgAspMetThrValProThrAlaMetTrp 352
   |||||
1013  CGCCTCCCTACTATGATGTGTGAGCATGACGCTGCCAATTCAGTGTGG 1062
353  ThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLe 369
   |||||
1063  AACGTTGGCCATGATCTCTGGCTGATCCCCAAGATGCGCAATCTGCT 1112
369  uSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAlaH 386
   |||||
1113  TCCCAAACTCCCAACCTTCTGTACCATAGGAGATTTCTCCCTACAATC 1162
386  lsValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGlu 402
   |||||
1163  ACCTGGACTTCATCTGGGCGATGGTGGCTCAAGAGGTTTACAATGAG 1212
403  IleIleHisLeuMetGlnGlnGluThrAsn 413
   |||||
1213  ATAGTTACCATGGGAGAGACTTAACAGAAAT 1245
seq_name: gb_hlc:AK009300

```

```

seq_documentation_block:
LOCUS      AK009300      1349 bp      mRNA      HTC      08-FEB-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:2310011G18, full insert sequence.
ACCESSION      AK009300
VERSION      AK009300.1  GI:12844011
KEYWORDS      CAF trapper.
SOURCE      Mus musculus
            (strain:C57BL/6J) adult male tongue cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:2310011G18.
ORGANISM      Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

```



```

189 hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
|||||  |||||||  |||||||  |||  |||||||  |||  |||||||  |||
573 CCACATCGGTTTATTCGCTTTTCTACCAATCTGCTGGCTAAACAA 622

206 IleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLys 222
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||
623 ATCAAGAGGTTTATGCAATAGCTCCAGTTGCTACTGTGAAGTATACAGA 672

222 sSerProGlyThrLysPheLeuLeuProAspMetMetIleLysGlyL 239
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
673 AAGTCCCTTAAAGATTTTCACTTATCTTCAAGTTCTTCTCAAGGTGA 722

239 eupheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu 252
|||||  |||||||  |||  |||  |||  |||  |||  |||  |||  |||
723 TATTGTGTAACAAGTTGTCATAGCTCCACCACTACTAGTACCAATTTCT 772

253 ArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleCy 269
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
773 GTACGGAAGTG.....TGCTACGGGAGCTGCTAGATCTTCTCTG 813

269 sSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnMetAsnM 286
|||||  |||||||  |||  |||  |||  |||  |||  |||  |||  |||
814 CAGCAACGCTTATTCATCTCTCTGCTGATTTGACAGAAACCTTAATG 863

286 etSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
|||||  |||||||  |||  |||  |||  |||  |||  |||  |||  |||
864 TGATCGCTTGTGTGTATCTAGGCAATATCCAGCAGCAATCTTACT 913

303 GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
|||||  |||||||  |||  |||  |||  |||  |||  |||  |||  |||
914 CAGACCTTTTCCACTGGGACAGCTTGTCTAAATCTGGGAGCTTCAAGC 963

319 aPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnProt 336
|||||  |||||||  |||  |||  |||  |||  |||  |||  |||  |||
964 CTATACTGGGAGTCCATTACAGACATGTTTACACTACATCAGAA 1013

336 hrProValArgTyrArgValArgAspMetThrValProThrAlaMetTrp 352
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1014 CGCTCCCTACTATGATGTGTCAGCCATGACCGTGCCAAATTCAGTGTG 1063

353 ThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLe 369
|||||  |||||||  |||  |||  |||  |||  |||  |||  |||  |||
1064 AACGCTGGCATGACATCCGCTGCTGATGCCCAAGATGTCGAATGCTCT 1113

369 uSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAlaH 386
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1114 TCCCAAACTCCCAACCTTCTGTACCATAGGAGATCTTCCCTACAAATC 1163

386 lsValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGlu 402
|||||  |||||||  |||  |||  |||  |||  |||  |||  |||  |||
1164 ACCTGGACTTCATCTGGCGGATGATGCGCTCAAGAGGTTTCAATGAG 1213

403 IleIleHisLeuMetGlnGlnGluThrAsn 413
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1214 ATATTACCATGATGGCAGAGACTAACAGAT 1246

seq_name: gb_htc:AK009474
seq_documentation_block:
LOCUS AK009474 1349 bp mRNA HTC 08-FEB-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:2310022103, full insert sequence.
ACCESSION AK009474
VERSION AK009474.1 GI:12844295
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:2310022103.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)

```

Carrinci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Methods Enzymol. 303, 19-44 (1999)
2 (sites)
Carrinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
3 (sites)
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carrinci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,
Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, F.,
Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T.,
Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsuura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
4 (sites)
The RIKEN Genome Exploration Research Group Phase II team and
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 (bases 1 to 1349)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Carrinci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,
Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiroka, T., Hori, F.,
Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J.,
Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T.,
Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y.,
Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H.,
Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, Y., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A.,
Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5'-GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTNN 3']. cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0 and subtraction to
Rot = 25.0. Second strand cDNA was prepared with the primer adapter
of sequence [5'-
GAGAGAGAGATCTCGAGTTTAATTAATTAATTCCTCCCCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

FEATURES
source

Location/Qualifiers
1. .1349
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGI:1898389"
/db_xref="MGI:1914967"
/clone="2310022103"
/sex="male"
/tissue_type="tongue"

/cqn2_6/ptodata/1/ina/5B_COMB.seq:US-08-600-908A-11 +	82.50	140.26	3.39
/cqn2_6/ptodata/1/ina/6A_COMB.seq:US-08-683-838A-11 +	82.50	140.26	3.39
/cqn2_6/ptodata/1/ina/5B_COMB.seq:US-08-483-101-1 +	82.50	122.45	33.26
/cqn2_6/ptodata/1/ina/5A_COMB.seq:US-08-348-891A-1 +	82.50	109.43	176.52
/cqn2_6/ptodata/1/ina/5A_COMB.seq:US-08-905-817-1 +	82.50	109.43	176.52

seq_name: /cqn2_6/ptodata/1/1na/5A_COMB.seq:US-08-227-108-2

seq_documentation_block:
: Sequence 2, Application US/08227108
: Patent No. 5807/26

STREET: 1335 AVENUE OF THE AMERICAS
CITY: New York

CITY: NEW YORK
STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible

```

;
; OPERATING SYSTEM:  PC-DOS/MS-DOS
;
; SOFTWARE:  Datont In Release #1.0
; Version #1.25

```

APPLICATION NUMBER: US/08/227,108

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 312 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1137 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1137
US-08-227-108-2

alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
US-09-333-159-47 x US-08-227-108-2 ..
Align seg 1/1 to: US-08-227-108-2 from: 1 to: 1137

37 ProThrLysAlaValaspProGluAlaPheMetAsnIleSerGluIle1 53
||||| :|||||: -|||||:|||||:|||||:|||||:|||||:
19 CCCACA.....AACCCCTGAAGTCACCATGAATATAGTCAGATGAT 59

53 eGlnHisGlnGlyTyProCysGluGluTyTGluValAlaThrGluAspG 70
| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
60 CACCTACTGGGATACCCACTGAGGAATATGAAGTGTGTGACCGAAGACG 109

70 lYtRileuSerValAsnArgIleProArqGlyLeuValGlnProLys 86

```


Tue Sep 4 10:56:04 2001

...CAGGAAATACCTGGTCTCGGAACAATFRO

402 GAGGAATAACCTGGCTCGCGAACAATAGACACTCTCAGTTTCTCAGGAT 451

150 GluPheTrfAlaPheSerTyrAspGluMetAlaArgPheAspLeuProAl 166
|||||
452 GAATCTGGCGCTTTCACCTATGATGAGATGGCAAAATATGACCTACCAG 501

166 aValIleAsnPheIleLeuGlnLysThrGlyGlnGluLysIleTyrTrv 183
|||||
502 TTCCATTAPCTTCATTCTGAATAAACTGGCCCAAGAACAGTGTATTATG 551

183 alGlyTyrSerGlnGlyThrThrMetGlyPheIleAlaPheSerThrMet 199
|||||
552 TGGGTCAATCTCAAGGCACACATATAGGTTTTATAGCATTTTCACAGATC 601

200 ProGluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaProIleAl 216
|||||
602 CTGAGCTGGCTAAAGAGATTAAATGTTTTTTGGCCCTGGGTCTGTGGC 651

216 aThrValLysHisAlaLysSerProGlyThrLysPheLeuLeuLeuProA 233
|||||
652 TTCCGCTGCCCTTCTGTACTAGCCCTATGGCCAAATTTAGGACGATTACCAG 701

233 spMetMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThr 249
|||||
702 ATCACTTCATTAAAGGACTATTATGGAGACAAAGAATTTCTCCCGAGAT 751

250 ArgPheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAs 266
|||||
752 GCGTTTTTGAAGTGGCTGGGTACCCAGCTTTGGCACTCATGTACTAGAA 801

266 pGlnIleCysSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnA 283
|||||
802 GGAGCTCTGTGGAAATCTCTGTTTTCTCTGTCTGGATTAAATGAGAGAA 851

283 snMetAsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGly 299
|||||
852 ATTTAAATATGCTAGAGTGGATGTATATACACACATTCTCCTGCTGGA 901

300 ThrSerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGl 316
|||||
902 ACTTCTGTGCAAAACATGTTACACTGGAGCCAGCGTGTAAATTCACAAA 951

316 uLeuArgAlaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysA 333
|||||
952 GTTTCAGAGCTTGTACTGGGAGCAGCTGCCAAGAATTTATTTTCATTACA 1001

333 snGlnProThrProValArgTyrArgValArgAspMetThrValProThr 349
|||||
1002 ACCAGAGTTATCCCCACATACAAATGTTGAAGGACATGCTTGCCCGACT 1051

350 AlaMetTrpThrGlyGlyGlnAspTrpLeuSerAsnProGluAspVally 366
|||||
1052 GCAGTCTGGACGGGGTCCAGACTGGCTTGCAGATGCTACGACGCTCAA 1101

366 sMetLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsnIleProG 383
|||||
1102 TATCTTACTGACTCAGATCACCAACTTGGTGTTCATCAGAGCATTCGCG 1151

383 luTrpAlaHisValAspPheIleTrpGlyLeuAspAlaProHisArgMet 399
|||||
1152 AATGGGAGATCTCTGACTTCATTATTTGGGGCCTGGATGCCCTTGGAGGCTT 1201

400 TyrAsnGluIleIleHisLeuMetGlnGln 409
|||||
1202 TATAATAAATTAATTAATCAATGAGGAAA 1231

seq_name: gb_pr9:HSLYACL

seq_documentation_block:

LOCUS	HSLYACL	2481 bp	mRNA	PRI	17-FEB-1997
DEFINITION	H.sapiens (HepG2) LAL mRNA for lysosomal acid lipase.				
ACCESSION	Z31690				
VERSION	Z31690.1	GI:506430			